REPORT RESUMES

ED 020 677 56 EM 006 299

A COMPARISON OF LINEAR AND BRANCHING TECHNIQUES OF PROGRAMED INSTRUCTION IN PLANE GEOMETRY. COMPARATIVE STUDIES OF PRINCIPLES FOR PROGRAMING MATHEMATICS IN AUTOMATED INSTRUCTION, TECHNICAL REPORT NO. 1.

BY- BEANE, DONALD G.
ILLINOIS UNIV., URBANA
REPORT NUMBER NDEA-7A-806 PUB DATE JUL 62
EDRS PRICE MF-$0.50 HC-$4.92 121P.

DESCRIPTORS- *PROGRAMED INSTRUCTION, *LINEAR PROGRAMING, *CONVENTIONAL INSTRUCTION, ABILITY, GEOMETRY, *BRANCHING, STUDENT ATTITUDES, *RETENTION, PROGRAMED MATERIALS, HENNON NELSON TEST OF MENTAL ABILITY, B.F. SKINNER, NORMAN CROWDER

SIXTY-FIVE STUDENTS IN TWO CLASSES IN HIGH SCHOOL GEOMETRY WERE ASSIGNED BY STRATIFIED RANDOM PROCEDURE ON THE BASIS OF THE HENNON-NELSON TEST OF MENTAL ABILITY TO FOUR EXPERIMENTAL GROUPS--TWO USING A LINEAR OR A BRANCHING TYPE PROGRAM EXCLUSIVELY, AND TWO SWITCHING PROGRAM TYPE MIDWAY THROUGH THE EXPERIMENT. A THIRD CLASS, TAUGHT BY THE SAME TEACHER, WAS GIVEN CONVENTIONAL INSTRUCTION. BOTH PROGRAMS WERE CONSTRUCTED FROM THE TEXT NORMALLY USED BY THE CLASS, THE LINEAR WITH VERY SMALL STEPS, AND THE AN ACHIEVEMENT TEST WAS ADMINISTERED AS PRETEST, POST-TEST, AND AN AHIEVEMENT TEST WAS ADMINISTERED AS PRETEST, POST-TEST, AND SEVEN WEEK DELAYED RETENTION TEST. AN ATTITUDE QUESTIONNAIRE WAS GIVEN HALFWAY THROUGH THE EXPERIMENT, WITH THE POST-TEST, AND WITH THE RETENTION TEST. ANALYSES OF VARIANCE SHOWED NO SIGNIFICANT DIFFERENCES AMONG EXPERIMENTAL GROUPS IN ACHIEVEMENT OR RETENTION, ALTHOUGH A SUB-GROUP OF HIGH ABILITY PERFORMED SIGNIFICANTLY BETTER THAN ONE OF LOW ABILITY STUDENTS IN ALL CONDITIONS. ALTHOUGH ALL EXPERIMENTAL GROUPS PREFERRED PROGRAMED TO CONVENTIONAL INSTRUCTION DURING AND IMMEDIATELY AFTER THE EXPERIMENT, NO PREFERENCE WAS SHOWN SEVEN WEEKS LATER. THE LINEAR PROGRAM WAS PREFERRED TO THE BRANCHING PROGRAM BY THE ENTIRE GROUP, ALTHOUGH IT TOOK SIGNIFICANTLY LONGER TO COMPLETE THAN THE BRANCHING PROGRAM. (BB)

ED020677

A-806
PA56

UNIVERSITY OF ILLINOIS

Urbana, Illinois

A COMPARISON OF LINEAR AND BRANCHING TECHNIQUES OF PROGRAMED INSTRUCTION IN PLANE GEOMETRY

DONALD G. BEANE

A
EM 006299

Technical Report No. 1

July, 1962

COMPARATIVE STUDIES OF PRINCIPLES FOR PROGRAMING MATHEMATICS IN AUTOMATED INSTRUCTION

Co-Investigators:

Lawrence M. Stolurow
  Professor, Department of Psychology
  Training Research Laboratory

Max Beberman
  Professor, College of Education
  University of Illinois Committee
  on School Mathematics (UICSM)

Project Sponsor:

Educational Media Branch
  U. S. Office of Education
  Title VII

Project No. 711151.01

UNIVERSITY OF ILLINOIS
Urbana, Illinois

A COMPARISON OF LINEAR AND BRANCHING TECHNIQUES OF PROGRAMED INSTRUCTION IN PLANE GEOMETRY

DONALD G. BEANE

Technical Report No. 1.
July, 1962

COMPARATIVE STUDIES OF PRINCIPLES FOR PROGRAMING MATHEMATICS IN AUTOMATED INSTRUCTION

Co-Investigators:

Lawrence M. Stolurow
  Professor, Department of Psychology
  Training Research Laboratory

Max Beberman
  Professor, College of Education
  University of Illinois Committee
  on School Mathematics (UICSM)

Project Sponsor:

Educational Media Branch
  U. S. Office of Education
  Title VII

Project No. 711151.01

EM 006 299

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ACKNOWLEDGEMENTS

The author wishes to express his special thanks to Dr. K. B. Henderson, Chairman of the Committee, and to Dr. L. M. Stolurow, both of whom gave generously of their time and assistance in carrying the study through to its completion. The helpful suggestions given by the other members of the Committee, Dr. R. E. Pingry, Dr. M. Beberman and Dr. W. M. Zaring, are also appreciated.

Special mention should be made of the help given by George Dodson of the Statistical Service Unit in processing the IBM cards and thanks given to Mrs. Julie Watson and Mrs. Elaine Wilson for their work in typing and reproducing the materials used in the study.

Appreciation is expressed to the faculty members and students of Rantoul Township High School and Urbana High School who cooperated so willingly in the data collecting process. Grateful acknowledgement is made of the financial aid given through a grant from the U. S. Office of Education, Title VII, Project No. 711151.01, which helped make the study possible.

Finally, sincere appreciation is given to Marjorie Beane for help and inspiration through all phases of the project.

TABLE OF CONTENTS

Page

LIST OF TABLES vi

LIST OF FIGURES vii

CHAPTER I. THEORETICAL BACKGROUND OF THE PROBLEM 1

Comparison of the Two Basic Positions 2

Research Hypotheses 6

CHAPTER II. RELATED RESEARCH 9

Effects of Response Modes, Step Size and Step Sequence on Achievement and Retention 9

Interaction of Ability and Step Size on Achievement 13

Effect of Programed Instruction on Attitudes 15

Programed versus Conventional Textbook Instruction 19

Summary 19

CHAPTER III. METHOD 21

The Sample 21

Procedure 22

Materials 26

CHAPTER IV. RESULTS 50

Analysis of Data on the Independent Variables 50

Analysis of Data on Achievement 53

Analysis of Data on Retention 56

Summary of Data on Attitude Toward Programed Instruction . 60

Analysis of Attitude Data 64

Summary of Data Indicating a Preference for the Linear or Branch Program 64

Page

Analysis of Program Errors 68

Time and Efficiency of the Experimental Treatments 72

Control Group Data 75

CHAPTER V. CONCLUSIONS AND SUMMARY 84

Conclusions Regarding Tests of the Hypotheses 84

Comparison of These Findings with Related Studies 88

Summary 92

BIBLIOGRAPHY 96

APPENDIX A. CRITERION ACHIEVEMENT TEST 98

APPENDIX B. ATTITUDE QUESTIONNAIRE 105

APPENDIX C. TIME SHEET 108

APPENDIX D. DATA ON INDIVIDUAL SUBJECTS 109

APPENDIX E. DISTRIBUTION OF RESPONSES TO INDIVIDUAL ITEMS ON THE ATTITUDE QUESTIONNAIRE (SECTION 1) 114

VITA 116

LIST OF TABLES

Table Page

1. Concepts and Principles 30

2. Symbol Combinations with Examples 33

3. Geometric Figural Context 34

4. Sequence of Ordered Triples (Concepts, Symbols, Figure) Included in the Linear and Branch Programs 36

5. Mean Scores of Treatment Groups and Ability Subgroups on the Variables of IQ Scores, Algebra Grades and Geometry Grades 51

6. Analysis of Variance Summary for the Independent Variables 54

7. Mean Scores on Achievement Post-Test 55

8. Analysis of Variance of Achievement Scores 55

9. Mean Scores on Delayed Achievement Test 56

10. Analysis of Variance of Delayed Achievement Scores 57

11. Mean Retention Scores 58

12. Analysis of Variance of Retention Scores 59

13. Mean Scores of Attitude Toward Programed Instruction . . . 61

14. Analysis of Variance of Attitude Scores 65

15. Mean Preference Scores Comparing the Linear and Branch Programs 66

16. Significance of Differences from Hypothesized Mean of 20 for Preference Scores of Combined Mixed-Treatment Groups and Ability Subgroups 68

17. Distribution of Errors for Individual Questions in the Linear and Branch Programs 69

18. Mean Error Rates (in per cent) of Treatment Groups and Ability Subgroups on the Linear and Branch Programs 71

Table Page

19. Mean Time (in minutes) Spent on the Programed Materials 72

20. Analysis of Variance of Times Spent on the Programed Materials 73

21. Mean Efficiency Scores of the Treatment Groups and Ability Subgroups 74

22. Analysis of Variance of Efficiency Scores 75

23. Means of Control Group and Experimental Groups on the Independent Variables 77

24. Analysis of Variance Comparing the Control and Experimental Groups on the Independent Variables 78

25. Means of the Control Group and the Experimental Groups on the Dependent Variables 80

26. Analysis of Variance Comparing the Control and Experimental Groups on the Dependent Variables 81

LIST OF FIGURES

Figure Page

1. Mean Attitude Scores Comparing Programed Instruction with Traditional Classroom Instruction 63

2. Mean Preference Scores for the Linear or Branch Program at the End of the Experiment 67

CHAPTER I

THEORETICAL BACKGROUND OF THE PROBLEM

The rapidly growing interest in automated instructional materials brings into focus the need for systematic research to evaluate various programing techniques. The programed materials in the current teaching machines and programed texts ordinarily present information to the student in the form of completion or multiple-choice questions. The student responds to the question, and then the answer is revealed. He next moves to further information which builds upon the preceding knowledge. This procedure implements certain psychological principles to facilitate learning. These are the following: 1) The learning task is organized into small, sequential steps. 2) The student is kept active by continually responding to the information contained in the program. 3) He can proceed at his own rate. 4) He gets continual feedback by being told if each response is right or wrong.

While there is general agreement about the psychological basis of programing, considerable disagreement exists as to what specific programing techniques can produce the desired learning most effectively. One group asserts that learning is facilitated with a small step, linear program by keeping errors to a minimum so that correct responses will be reinforced. Another group asserts that learning is facilitated through reinforcement in a branch program by including explanatory material when errors occur. Two basic positions relative to programing techniques have evolved. One position has been advanced by the psychologist, B. F. Skinner of Harvard University. The other position has been championed by another psychologist, Norman A. Crowder.

Professor Skinner is primarily responsible for the revival of interest in teaching machines. Furthermore, he is responsible for suggesting programed instruction in the Harvard Educational Review in 1954. (18)[1] It seems that the great advances in industrial automation achieved by the 1950's provided a receptive audience for Skinner's suggestions on an alternative to current educational practices. His proposal that automation could revolutionize education as it had industry was met with enthusiasm by many.

Comparison of the Two Basic Positions

Skinner's method as outlined in his article in Science (19), has been related to his notions about conditioning. However, in programed instruction Skinner and others are interested in shaping and maintaining desirable forms of verbal behavior in humans through immediate positive reinforcement. This differs from operant conditioning in that the response when reinforced results in the stimulus being changed. Skinner has described certain programing techniques which he believes will aid in the effective shaping and maintaining of the desired verbal behavior. One technique is that of writing frames in such a way that the student must construct his own response. Completion types of questions are used for this. States Dr. Skinner,

---

[1] In referring to references, two figures will be used. The first one denotes the number of the reference as it is placed in the bibliography. The second figure, when used, designates the page or pages from which the quotation has been made.

> "An appropriate teaching machine will have several important features. The student must compose his response rather than select it from a set of alternatives, as in a multiple-choice self-rater. One reason for this is that we want him to recall rather than recognize - to make a response as well as see that it is right. Another reason is that effective multiple-choice material must contain plausible wrong responses, which are out of place in the delicate process of 'shaping' behavior because they strengthen unwanted forms." (19:970)[2]

Secondly, the program should be constructed with very small steps so that the correct answer is almost automatic and very few errors are made. Skinner states his case for keeping programs very easy in the following terms:

> "Can material be too easy? The traditional teacher may view these programs with concern. He may be particularly alarmed by the effort to maximize success and minimize failure. He has found that students do not pay attention unless they are worried about the consequences of their work. The customary procedure has been to maintain the necessary anxiety by inducing errors. In recitation, the student who obviously knows the answer is not too often asked; a test item which is correctly answered by everyone is discarded as nondiscriminating; problems at the end of a section in a textbook in mathematics generally include one or two very difficult items and so on . . . Making sure that the student knows he doesn't know is a technique concerned with motivation, not with the learning process. Machines solve the problem of motivation in other ways. There is no evidence that what is easily learned is more readily forgotten. If this should prove to be the case, retention may be guaranteed by subsequent material constructed for an equally painless review." (19:975)

Thirdly, Dr. Skinner favors a linear program in which all the students are given the same sequence of steps rather than a branch program. In advising the programer about the procedure to use in composing a program, he states,

---

[2] Empirical data exist which indicate that "composing the response" may not be a critical factor. See Walker (22).

> "A first step is to define the field. A second is to collect technical terms, facts, laws, principles, and cases. These must then be arranged in a plausible developmental order - linear if possible, branching if necessary." (19:974)

Skinner also suggests,

> "A program designed for the slowest student in the school system will probably not seriously delay the fast student, who will be free to progress at his own speed. (He may profit from the full coverage by filling in unsuspected gaps in his repertoire.)" (19:976)

An example of linear programing is Holland and Skinner's text, The Analysis of Behavior (13).

A different position in regard to effective programing techniques has been taken by the psychologist, Norman Crowder. He considers the problem of programed instruction as one of communication between programer and learner rather than one of conditioning the learner to make the proper verbal response. He favors multiple-choice questions with larger step size rather than the small step advocated by Skinner. Also, he favors a branch program rather than a linear, with the student's choice of an answer as the determining factor in the sequence of material. Crowder presents his position in the following terms:

> "'Automatic tutoring' is an individually used, instructorless method of teaching which represents an automation of the classical process of individual tutoring. The student is given the material to be learned in small logical units (usually a paragraph or less in length) and is tested on each unit immediately. The test result is used automatically to control the material that the student sees next. If the student passes the test question, he is automatically given the next unit of information and the next question. If he fails the test question, the preceding unit of information is reviewed, the nature of his error is explained to him, and he is retested. The test questions are multiple-choice questions and there is a separate set of correctional materials for each wrong answer that is included in the multiple-choice alternative. The technique of using a student's choice of an answer to a multiple-choice question to determine the next material to which he will be exposed has been called 'intrinsic programming'." (5:40)

Crowder has proposed several specific devices for presenting intrinsic programs. However, the simplest device is a specially prepared book in which each alternative answer is identified with a page number. The page numbers in the book are assigned randomly so that the reader cannot progress from one page to the next except by actively responding to each question. These books have been referred to as "scrambled texts"; an example is Crowder and Martin's text, Adventures in Algebra (7). The variety of types of programs that can be prepared in intrinsically programed form is limited only by the ingenuity of the programer. The simplest form of program step is one in which each wrong answer refers the student (after further discussion) back to the original choice page to try again.

Crowder defends his belief in flexible program formats and flexible step size as follows:

> "To sum up what has been said, we approach the design of a teaching machine as a problem in communication. The conditions of the program are such that the greatest flexibility, both within and between program steps, is required. The within-steps flexibility is required because we wish to communicate complex information to a complex organism, that is, an intelligent human being. To provide the necessary flexibility within steps the devices provide that any unit of information up to the amount that can be presented on a single page or a page size viewing screen can be presented at a single presentation. The requirement for flexibility between steps arises because communication may fail, particularly if we are attempting to get the student to flex his mental muscles a little and give him fairly stiff questions." (5:52).

The two theoretical positions of Skinner and Crowder present contrasting views on three experimentally manipulative variables. These variables are 1) size of step, 2) form of step, and 3) sequence of steps. Skinner's position favors small steps with completion type questions in which students construct the answer. The order of steps is linear in which all students follow the same sequence. Crowder's position favors the use of larger steps in the form of multiple-choice questions. The sequence to be followed depends upon the student's answers, and therefore a branching program is needed.

The purpose of this study is to gather evidence on the relative effectiveness of these two basic types of programs using subject matter from high school plane geometry. Specifically, do these two basic types of programs produce significantly different results in achievement, retention or attitude toward programed instruction when used by students of different ability levels?

Research Hypotheses

Considerable evidence has been reported that students of a wide range of ability can learn material from various school subjects when either a linear or a branch program is used.[3] This evidence gave rise to the following hypotheses:

1. No differences in achievement in plane geometry will exist between groups of students using a linear or a branching program.
2. No differences in achievement in plane geometry will exist between high and low ability students following programed instruction.

Research evidence is scarce on the variable of retention following programed instruction. Since no conclusive evidence exists stating that one type of program is superior for retention, the following hypotheses are made:

---

[3] This evidence will be discussed in Chapter II: "Related Research".

3. There will be no differences in retention between the treatment groups.

4. There will be no differences in retention between the high and low ability levels.

The next variable to consider is attitude toward programed instruction. The branching program was deliberately made more difficult by including more material in each frame. The belief was that this would make the program more interesting and challenging to the better students. In contrast, the linear program was deliberately made easy by constructing very small steps. This was done to adhere to Skinner's theory that all students could experience a high degree of success using a small step program and this positive reinforcement would be self-motivating. The prediction was made that the high ability students would indicate a significant preference for the more difficult branch program, and the low ability students would show a significant preference for the easier linear program. However, the hypotheses are stated in the null form to test for a significant preference for either program.

5. The high ability students will show no preference for either the linear or the branch program.

6. The low ability students will show no preference for either the linear or the branch program.

In addition to these explicitly formulated hypotheses the study was designed to obtain information on general attitude toward programed instruction in geometry as opposed to classroom instruction with a teacher and regular textbook. Also, the time spent on the programed materials would be compared with achievement to obtain a measure of the relative efficiency of the programs.

The decision was made to have a teacher-taught control group for the purpose of judging the effectiveness of the experimental treatments on achievement and retention against a group using regular classroom instruction with teacher and textbook. This comparison was supplementary to the main comparisons between the experimental, programed instruction groups. Nevertheless, the validity of any experimental teaching method should be judged not only on its relative effectiveness in comparison with other experimental methods, but also in comparison with established methods of teaching. The latter is the more difficult to accomplish in a meaningful way. The difficulty arises from the lack of information about the sampling distribution of the sets of methods that are compared. Thus in any specific comparison one does not know whether each method is above, below or average within its own set.

In this study the following hypothesis is tested:

7. There will be no differences between the control group and the experimental groups in achievement or retention.

CHAPTER II

RELATED RESEARCH

Programed instruction is a relatively new field of research in education. Lumsdaine and Glaser (15) and Stolurow (21) give a comprehensive treatment of the theoretical positions in programing and the preliminary research that has been done in the field up to 1960. However, the field is developing rapidly. More research is presently being conducted in programed instruction than has been reported up to this time in the literature.

This review of related research includes those studies which compared at least one of the program variables - response mode, step size or step sequence - which are investigated in the present study. In addition, a few studies have been included which measured attitudes or which compared programed instruction and conventional text material.

Effects of Response Modes, Step Size and Step Sequence on Achievement and Retention

Coulson and Silberman (4) investigated the effects of three variables in programing a portion of a college course in elementary psychology used at Harvard University. The experimental variables, each having two possible values, are as follows: a) response mode (multiple-choice versus constructed response), b) size of steps between successive items to be taught (small steps versus large steps) and c) type of step sequence control (branching versus nonbranching). The eight combinations of these three variables constituted the eight treatments compared in the experiment. Each treatment group contained ten students.

A 36 question criterion test, of which 19 were constructed answer and 17 multiple-choice, was given after the experiment to measure achievement and again three weeks later to measure retention. Approximately two hours total time was spent on the programed materials. The dependent variables were the required teaching machine training time and scores on the criterion test. The results of this experiment were as follows:

1) No significant differences were obtained among the eight experimental groups on the total criterion test. However, the differences in time taken to complete the program were significant. The constructed-response training condition took more time than the multiple-choice condition; the small step condition took more time than the large step condition and the nonbranching condition took more time than the branching condition.

2) No significant difference in retention was found when the mean of the first administration of the criterion test was compared with the mean of the second administration three weeks later.

In evaluating the results, it must be kept in mind that only two hours were spent on the programed materials. Perhaps longer exposure would produce significant differences between treatment groups in achievement.

Fry (11) investigated the relative effectiveness of two response modes on achievement and retention. Sixteen Spanish words were taught to ninth grade English speaking students using a teaching machine device which could be programed to handle either multiple-choice questions or questions requiring the construction of an answer. Each multiple-choice question contained four alternatives.

Three conditions were compared in this study. In condition I, students worked to the same criterion of mastery: two correct responses to each of the 16 items. In condition II, equal total working time was controlled by stopping all students prior to completion of the learning task by the fastest student. In condition III, time and number of repetitions were controlled by presenting the stimulus material on a large flash card to both response groups simultaneously. A post-test was given immediately after training and a delayed test two days later. The test consisted of the 16 training items; eight were constructed items and eight multiple-choice items. On the delayed test the format was reversed with constructed item words appearing in multiple-choice form and vice versa. The number of subjects used in each condition was: I - 81, II - 66 and III - 153.

The results indicated that under all three conditions, the constructed response group did significantly better than the multiple-choice response group on the constructed response sub-test of the immediate test and the delayed test. No significant differences existed between the two groups under any of the three conditions on the multiple-choice sub-test of either the immediate test or the delayed test. However, both groups averaged more than seven out of eight possible items correct on the multiple-choice sub-test under all three conditions on both tests. A reasonable explanation is that the multiple-choice sub-test was too easy to discriminate differences if they did exist.

The total working time was recorded for each student in condition I. The mean total training time for the constructed-response group was 14.2 minutes and for the multiple-choice group was 8.3 minutes. The difference was significant at the .01 level.

Evans, Glaser and Homme (10) at the University of Pittsburgh conducted two preliminary studies using a programed textbook format. They constructed a program designed to teach "conversion to number bases other than ten". Then they deleted or added items to produce four versions of the following lengths: 30, 40, 51 and 68 steps. Four independent groups of five graduate students in education were given these sequences. Each student took a written test after he completed the sequence. The results indicated that "smaller" steps, i.e., the use of more frames to cover the same subject matter, were associated with significantly fewer errors on immediate and delayed written tests. Also, smaller steps were associated with fewer errors during learning. There is probably a point of diminishing returns in decreasing the step size since the scores of group D (68-step sequence) were slightly lower than those of group C (51 steps).

Silberman, Melaragno, Coulson and Estavan (17) conducted an experiment in which a computer-controlled teaching machine was used to evaluate the effectiveness of adapting sequences of material to the number of errors made on a particular topic. The subjects were 36 students who had just graduated or were in their last year of high school in Los Angeles County. Subjects were pretested with the Henmon-Nelson Test of Mental Ability and then randomly assigned to two groups (branching versus fixed sequence). The materials were 411 multiple-choice steps on logic which were put onto 35-millimeter slides for use in a random-access slide projector. A high speed, general purpose digital computer selected the step sequence, and the subjects used an electric typewriter to record their answers.

The branching group received sequences of questions determined by the errors made in the teaching session. The machine selected an approximate sequence of instructional material for each student based on his errors. The sequences were on four levels of difficulty varying in step size but covering the same material. A student with too high an error rate on one sequence would be given another sequence over the same material but on a lower difficulty level. Each member of the branching group was paired at random with one member of the fixed-sequence group and given the identical step sequences. The difference between the two groups was that the machine was responsive to the errors made by subjects in the branching group but not to errors made by subjects in the nonbranching group. The criterion test was composed of 51 multiple-choice and 44 free response items. Half of the test items were similar to actual training and half required application to new situations.

The two groups were compared on number of errors, training time in minutes, Henmon-Nelson IQ scores and post-training criterion scores. A covariance analysis of criterion scores using aptitude and training time as control variables yielded no significant differences between the branching and fixed sequence conditions. However, the authors report that an analysis of the criterion test showed that low aptitude subjects (Henmon-Nelson IQ Scores) in both groups failed to learn much of the material. Eleven of 13 low aptitude students in the two groups fell well below their group means on the criterion test. The authors propose that factors other than error rate may be more appropriate, or should be considered in addition to errors, in making branching decisions to accommodate individual differences.

Interaction of Ability and Step Size on Achievement

Shay (16) investigated the null hypothesis that there is no relationship between intelligence and step size on a teaching machine program for each of the criteria of: total learning, learning of 'rote' materials, learning of materials involving 'understanding', percentage errors on the program and time to complete the program. Size of step was defined as the "difficulty of giving the correct answer" and was inferred from the proportion of errors made on the program. Three programs were written covering a fourth grade unit about symbols and principles for construction of Roman numerals. The final experimental programs contained 103 (large step), 150 (medium step) and 199 (small step) frames.

Ninety subjects were selected on the basis of a separate Roman numeral pretest and were divided into three ability levels on the basis of a group intelligence test. The subjects in each ability level were assigned randomly to one of the three programs to form nine experimental groups of ten subjects each. A covariance analysis of criterion scores with the Roman numerals pretest as a control was used with the results indicated in the following table.

Shay was interested in the interaction between intelligence and treatment on the dependent variables. He found no significant interaction except on percentage of errors on the programs. He draws the conclusion that the data support Skinner's position that it is not necessary to provide more than one program on the basis of differential initial ability.

Interesting observations about main effects can be made from the data in the table. The type of program used had no significant effect upon scores on the total criterion test or its two parts. However, highly significant differences existed between ability groups on the total criterion achievement test and its parts. Other data in the report indicated that the above average ability group did better on achievement than the average group, which in turn did better than the below average group. Also, the above average group made fewer errors on the program than the other two groups and took less time to complete the programs. This is in contrast to studies cited by Stolurow (21) in which ability differences correlate almost zero with gain or post-test scores provided all students meet a minimum criterion of achievement.

		F				
Source	df	Total Posttest	'Understanding' Subtest	'Rote' Subtest	% Error	Time on Program
Programs	2	1.80	1.45	1.30		27.21**
Ability	2	10.43**	8.23**	7.78**		19.17**
Interaction	4	1.87	1.35	2.26	3.17*	.22
Within cells	80					

* .05 level of significance
** .01 level of significance

Effect of Programed Instruction on Attitudes

Cassel and Ullom (2) (3) conducted two studies to evaluate programed instruction of a course in computer mathematics using the Auto Tutor Mark II machine developed by the Western Design and Electronics Co., Goleta, California. The branching program techniques described by Crowder (5) were used to prepare the materials, which were a revised and extended machine adaptation of Crowder's text, The Arithmetic of Computers. (6)

The first study (2) involved 32 selected 9th and 12th grade students of high ability in the Lompoc Unified Schools, California. The total mental scale IQ on the California Test of Mental Maturity, Short Form (CTMM) and total score of the Iowa Tests of Educational Development (ITED) were averaged for all students having IQ's of 115 or higher in the 9th and 12th grades. The average scores were ranked and the top 16 boys and top 16 girls from each grade were invited to participate. All 64 students agreed to take part in the experiment.

Half of the participants, selected at random with equal numbers by sex and grade,were identified as the experimental group and assigned to the "teaching machine" course; the other half acted as a "no-instruction" control group. The control group was administered the criterion test, the Computer Mathematics Test, on two consecutive days to ascertain the effects of practice associated with two administrations of the criterion test. The experimental group was administered the criterion test before the experiment and as each student finished the course. A three way analysis of variance was made of gain scores resulting from pre- and post- administrations of the Computer Mathematics Test. The following comparisons between means were made: 1) Between Groups, 2) Between Grades, 3) Between Sexes, 4) Group X Grade, 5) Group X Sex, 6) Grade X Sex and 7) Group X Grade X Sex. The only F comparison which attained significance was that between the control and experimental groups which was significant beyond the .001 level. We must keep in mind in interpreting these results that the control received no instruction on the material covered in the test, so you would expect the experimental group to do significantly better. The authors report:

"a) The superiority of the experimental group over the control group, as measured by gains on the criterion test, was of both practical and statistical significance.

b) Gains on the criterion test were approximately equal for both sexes.

c) Gains on the criterion test were approximately equal for the two grades investigated." (2:225)

An attitude questionnaire was given to all students in the experimental group after one hour of machine instruction and at the end of the course. (The time spent on the course was not reported.) Percentages for each choice of answers were reported but no statistical analysis attempted because a number of students failed to return the questionnaires after the first hour of exposure to the programed materials. The conclusion was reached that generally the students strongly favored the use of the automated teaching technique. The real possibility that the novelty effect influenced attitudes favorably toward the machine instruction must be considered in interpreting the findings. Results on one of the attitude questions are presented above.

"If I had to study more of this kind of material I would prefer to use:"

	After 1 hour (%) N=19	At completion of course (%) N=31
Auto Tutor	78.95	80.65
Class Lecture	21.05	16.13
Typical Textbook	0.00	3.23

Cassel and Ullum (3) later conducted the same experiment with average ability students. Only five boys and five girls from both the 9th and 12th grades completed the course after 32 had been selected to participate.

A comparable control group was obtained and the same analysis of variance again disclosed that the only significant difference was between the programed instruction group and the "no instruction" control group.

The same attitude questionnaire was given to the average ability students and the results on the same question, reported above, are reported in the following table. A Chi-square test revealed that attitudes changed significantly on only one item out of the twelve in the questionnaire from the first administration to the second.

"If I had to study more of this kind of material I would prefer to use:"

	After 1 hour (%) N = 20	At completion of course (%) N = 20
Auto Tutor	90	80
Class Lecture	10	15
Typical Textbook	0	5

Skinner and Holland (20) investigated attitudes of students toward programed instruction in a college-level general education course in human behavior. They report:

> "Considering the fact that the student population was highly selected and contained many juniors and seniors of considerable college experience and high caliber, it appears to be encouraging that 99 percent felt that the machine helped them understand the text and that 78 percent felt that they learned more from the machine than from the text." (20:169)

Programed Versus Conventional Textbook Instruction

Evans, Glaser and Homme (10) evaluated programed learning and textbook presentation of the same material. In one study 17 undergraduates were given ten pages of a standard statistical text; a second group of 16 students were given a programed text covering the same material. After each group finished, they were given a multiple-choice test. The programed text group obtained higher mean performance scores but without significant differences. However, the programed text group showed significantly less variability in their scores.

Summary

Conclusive evidence exists that students can learn by programed instruction when the criterion is an achievement test. When the control groups receive conventional textbook instruction over the material covered on the criterion test, differences in achievement between the programed instruction groups and control groups are usually insignificant.

Conflicting evidence exists on the effects of ability differences on achievement following programed instruction. Also, the evidence comparing step size, step form or step sequencing is inconclusive.

Numerous studies indicate that a multiple-choice program takes considerably less time to cover the same material than a program in which responses must be constructed.

Students given an attitude questionnaire during or following programed instruction, in general, react very favorably to programed instruction. How much of this favorable attitude is due to the actual programed instruction and how much is due to the novelty of being in an experimental situation is difficult to determine.

CHAPTER III.

METHOD

The Sample

The sample for this study was taken from three plane geometry classes in Rantoul Township High School in Rantoul, Illinois. All three classes were taught by the same instructor. By examination of the teacher's lesson plan book and through personal interviews, it was determined that the teacher made a conscientious effort to present the same material to all three classes, to the point of giving identical assignments on the same day. Therefore, it could reasonably be assumed that the students in the three classes began the experiment with very similar instruction in plane geometry.

Two of the classes were designated as experimental classes and received the programed instruction covering a unit on parallel and perpendicular lines. The third class was designated as the control group and covered the same unit of material with the regular teacher. The instruction in the control group was a continuation of the same type of instruction all three classes had received prior to the experiment.

Design

The students in the experimental classes were divided randomly into four treatment groups. Group 1 used the branch program for the entire experiment. Group 4 used the linear program exclusively. Groups 2 and 3 used one program for the first half of the unit and then changed to the other program for the last half. The balanced design is indicated below, along with the added control group.

Group	First Half	Second Half
Experimental Group 1	Branch Program	Branch Program
Experimental Group 2	Branch Program	Linear Program
Experimental Group 3	Linear Program	Branch Program
Experimental Group 4	Linear Program	Linear Program
Control Group 5	Regular Instruction	Regular Instruction

Procedure

One week before the experiment began, all students in the three classes were administered the Henmon-Nelson Test of Mental Ability (Grades 9-12, Form A, 1957 Edition). Deviation IQ scores were obtained, using tables in the Examiner's Manual provided for this purpose. IQ scores of students in the two experimental classes were pooled and ranked. Students were assigned to one of the four experimental groups using a stratified random procedure. The ranked IQ scores were divided into sets of four each. A table of random numbers was used to assign the first three students in the top set of four to one of the experimental groups. The fourth person was assigned to the group not selected for the other three in the same set. In this way the top four students, based on their IQ scores, were each assigned to a different experimental group. The next four students were assigned to different experimental groups using the same procedure, and this process was continued until all students were assigned to one of the four experimental groups. The median IQ score was used to divide the students into a high and a low ability subgroup in each experimental group.

Pretesting

On the day before the experiment began, all students in the three classes were administered a criterion achievement test of the material to be covered in the experiment. This was done to determine among the groups their relative knowledge of the material to be covered by the various methods of instruction. See APPENDIX A for a copy of the criterion test. Students were given as much time as they wanted to work on the test and were asked to hand the test to the experimenter when they had finished.

Instructions

On the first day of the experiment the following instructions were read to the students by the experimenter:

> "The particular experiment in which you are participating is a comparison of programed instruction with regular classroom teaching."
>
> To control group: "This class will continue with regular classroom instruction under Mr. Coffey. However, your work will be compared with the two classes using programed materials. During this experiment I would like each of you to keep a record of how much time you spend outside of class working on your geometry assignments. This record will not affect your grade in geometry in any way. The person who indicates he is spending three hours every night on geometry will be considered no better nor worse than the person who finishes his assignments in class. Also, this experiment is no reason for spending a lot more or less time on geometry than you have been spending. We merely want to know how much time, on the average, students spend on their geometry assignments. I will collect these time sheets each Monday." See APPENDIX C for a copy of the time sheet.
>
> To experimental groups: "This class will be one of two classes using programed materials. We want to compare how well you can learn material presented in this manner in comparison with those learning the same material in a regular classroom situation. All your work will be done by yourself during class. No homework assignments will be given. This is really an experiment to see if you can teach yourself geometry. Two types of programs are being compared so you will not all be working with the same booklets."

The students were instructed to indicate on the answer sheet the time they started each day and the time they stopped. The general instructions printed in the front of the first booklet of each type of program were also read to the students by the experimenter.

Recording Student's Responses

Ball point pens were used exclusively during the experiment. This discouraged attempts to erase or change wrong answers. Emphasis was made of the fact that the students would be graded on their scores on the achievement test to be given at the end of the unit, not on the number of correct answers to questions in the programed materials. This was done in an effort to reduce the temptation to look ahead to find the right answer before indicating a choice on the answer sheet. Also, the experimenter observed the students at work during the entire class period each day to check that students followed instructions. All work of the experimental groups was done in class under the direct supervision of the experimenter. In case of absence a student was allowed to work after school under the supervision of the regular teacher.

Each student in an experimental group was given an answer sheet covering the questions in the first half of the programed unit. When a student finished the first half of the unit he turned in his answer sheet and was given a copy of the attitude questionnaire to complete. See APPENDIX B for a copy of the attitude questionnaire. Students were told to omit questions 8, 9 and 10 on the attitude questionnaire which dealt with a comparison of the two programs. When a student completed the questionnaire, he was given the answer sheet for the second half of the programed unit and the appropriate program booklet.

When a person finished the entire unit he was asked to complete the attitude questionnaire a second time. This time he was instructed to answer all the questions even though he had worked on only one type of program. If a student did not finish the unit in the allotted time of two weeks, he was asked to fill out the questionnaire on the last day of the experiment prior to the day the criterion achievement post-test was administered. All students were instructed to work through the review frames at the end of the unit even though they had not completed all the questions in the program. This allowed all students to benefit from the review. Although some students did not finish the programed materials in the time allotted, all students completed enough of the material to be able to answer the questions included in the criterion test.

Procedure Following Completion of Program

Those students who finished the programed materials early were given the first booklet of a programed unit, "Basic Concepts of Statistics". The attitude questionnaire was completed by these students who finished early before they were given the supplementary materials. This was done to insure that responses on the attitude questionnaire reflected attitudes about the programed materials presented as part of the experiment and not the supplementary programed materials.

Retention Testing

Seven weeks following the completion of the experiment, the attitude questionnaire and criterion test were administered a third time. The attitude questionnaire was administered in the first five minutes of the period and the rest of the period used for the criterion test. The students were given no advance notice that the criterion test or attitude questionnaire would be given a third time.

Materials

Content

The decision was made to program the material contained in the chapter, "Parallel and Perpendicular Lines", from the text, New Plane Geometry (24) by Welchons and Krickenberger. This traditional text has been very popular as evidenced by its eight copyright editions over the past 25 years. The purpose of this experiment was to obtain empirical data from geometry students of a wide range of ability using programed materials from a traditional text. This particular text was the one being used in the Rantoul Township High School. The axiomatic system presented in the text was used in the programed versions of the materials to maintain continuity with previous instruction. Also, the students would be returning to the text following the two-week experiment.

Studies involving programed versions of the UICSM materials and the SMSG materials are currently being conducted on a large scale. Both these experimental curricular programs stress set theoretical concepts more than the traditional geometry texts in use. Because of the studies being done with the nationally-known curricular programs, the decision was reached that this study could make a greater contribution by programing material from a traditional text.

The text used has certain drawbacks in terms of the mathematical content presented. Some of the mathematical inadequacies of the text, which are related to the material programed for this study, are given below.

1. Occasionally, weaknesses can be found in the rigor of the proofs. For example, in the exclusion method of indirect proof the assumption was made but not stated that one of the possible conclusions must be true. This assumption is necessary, because without it, eliminating all but one possible conclusion would not guarantee that the one remaining conclusion is true.

2. In a few instances, faulty concepts are presented. For example, *a theorem* is defined as a statement to be proved. Rather, it is a necessary conclusion that follows logically from the axiomatic system involved. Another example is the definition of *parallel lines* given in the text: "Two lines are parallel if they lie in the same plane and do not intersect even if extended." It is odd to speak of extending lines since they are infinite in length.

3. Occassionally, the notation is undesirable. For example, no distinction is made between an angle and the measure of an angle. Frequently statements are found in the text such as: "$\angle 3 = 30^{\circ}$". Mathematicians use the equals sign to designate that the two quantities represented are the same. However, an angle and a measure of 30° are not the same; therefore the equality symbol should not be used in this situation.

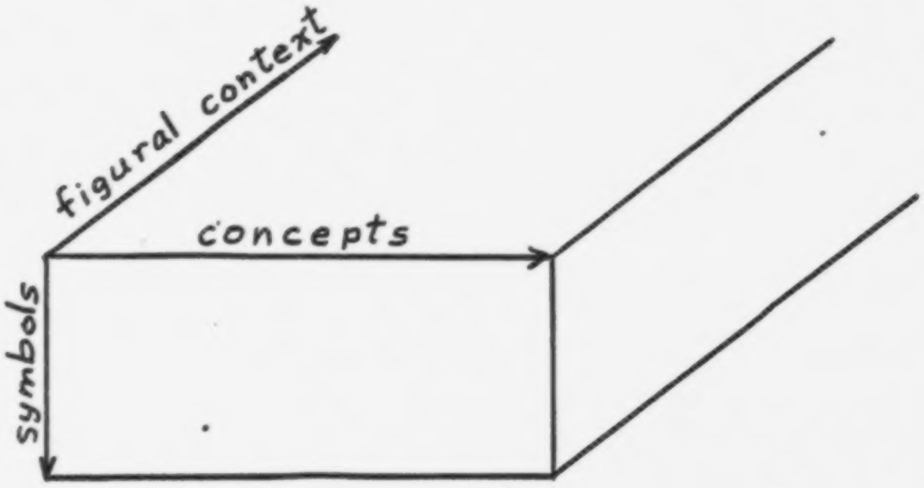
In a manner suggested by Stolurow,[4] a three dimensional representation of the content universe was used, from which a sample of items for the program was selected. One dimension represented the concepts and principles to be taught. A second dimension was the encoding--the combination of symbols--used to communicate the concepts and principles.

---

[4] Personal communication.

The third dimension was the figural context used to illustrate the concept or principle being taught.

The three dimensional universe is pictured below. Sample points included in the program can be represented as ordered triples (concept, symbols, figure).



Forty concepts and principles are presented in the program. Eight combinations of upper and lower case letters are used to represent lines and angles. Thirty-three different figures are used to illustrate the concepts and principles in the programs. These figures are contained in a separate 8-1/2"by 11" booklet. Space is provided for the students to write the steps of proofs in the Figures Booklet as they are presented in the program booklets. The forty concepts and principles are listed in Table 1. The eight combinations of symbols used are presented in Table 2. Verbal descriptions of the geometrical figures are given in Table 3. The programs present only a small sample of the possible ordered triples in the universe.

The sequence of the ordered triples (concept, symbols, figure) included in the programs is given in Table 4. The frame number for the linear and branch programs, listed after each ordered triple in Table 4, indicates when this combination first appears in each program.

The braces labelled "related problems" indicate applications of the concepts in the context of geometrical configurations different from those in which they were first presented in the program. This procedure is designed to help the student in generalizing the concepts and principles which are learned in a specific context.

Construction of the Linear Program[5]

The linear program contained 951 frames in five spiral bound booklets. The branch program contained 852 frames in seven booklets. The experiment was so designed that half of the students changed programs in the middle of the unit. The switch came in each program at the end of the same series of practical application problems.

The five booklets containing the linear program were 8-1/2"by 11" in size and each contained 50 pages with the exception of the last booklet which had 38 pages. The booklets were divided into sections of ten pages each with four frames on each page. The student worked through the ten frames on the top row of the pages in each section, then the second, third and fourth rows in that order. Every tenth page was printed on blue paper to serve as a cue to the student to return to the beginning of that section and answer the questions on the next row down. The students were instructed to write their answers to each question on a separate answer sheet and then to turn to the next page where the correct answer was printed to the left of the next question.

One of the psychological principles used in the construction of the linear program was Skinner's "vanishing" technique (19). Prompts

---

[5] Copies of the linear and branch programs are available for review from the University of Illinois Library, Urbana, Illinois. See Beane (1). Volume 1 contains the linear program and the supplementary Figures Booklet. Volume 2 contains the branch program.

Table 1

Concepts and Principles

---

1. Parallel lines

2. Postulate: Two straight lines in the same plane are either parallel or intersecting lines.

3. Transversal

4. Interior angles formed by a transversal intersecting two or more lines.

5. Exterior angles formed by a transversal intersecting two or more lines.

6. Alternate interior angles formed by a transversal intersecting two or more lines.

7. Alternate exterior angles formed by a transversal intersecting two or more lines.

8. Corresponding angles formed by a transversal intersecting two or more lines.

9. Contradiction Postulate: If a hypothetical statement leads to a contradiction of a known fact or hypothesis, the statement is false.

10. Exclusion method of indirect proof.

11. Theorem: If two lines form equal alternate interior angles with a transversal, the lines are parallel.

12. Corollary: If two lines form equal corresponding angles with a transversal, the lines are parallel.

13. Corollary: If two lines form supplementary interior angles on the same side of a transversal, the lines are parallel.

14. Corollary: Two lines perpendicular to a third line are parallel.

15. Parallel Postulate: Through a given point there can be one and only one parallel to a given line.

16. Corollary: Two lines parallel to a third line are parallel to each other.

17. Theorem: If two parallels are cut by a transversal, the alternate interior angles formed are equal.

Table 1

Concepts and Principles (con't)

---

18. Converse of a statement

19. Corollary: If two parallels are cut by a transversal, the corresponding angles formed are equal.

20. Corollary: If two parallels are cut by a transversal, the two interior angles on the same side of the transversal are supplementary.

21. Corollary: If a line is perpendicular to one of two parallel lines, it is perpendicular to the other.

22. Theorem: If two angles have their sides parallel, right side to right side and left side to left side, the angles are equal.

23. Theorem: The sum of the angles of a triangle is a straight angle or 180°.

24. Corollary: If two angles of one triangle are equal respectively to two angles of another triangle, the third angles are equal.

25. Corollary: A triangle can have no more than one right angle or one obtuse angle.

26. Corollary: An exterior angle of a triangle equals the sum of the two nonadjacent interior angles.

27. Corollary: The acute angles of a right triangle are complementary.

28. Corollary: If two right triangles have the hypotenuse and an acute angle of one equal respectively to the hypotenuse and an acute angle of the other, the triangles are congruent.

29. Theorem: If two angles have their sides perpendicular, right side to right side and left side to left side, the angles are equal.

30. Theorem: If two angles of a triangle are equal, the sides opposite these angles are equal.

31. Corollary: An equiangular triangle is equilateral.

32. Inverse of a statement

33. Contrapositive of a statement

34. Theorem: If two right triangles have the hypotenuse and a leg of one equal respectively to the hypotenuse and a leg of the other, the triangles are congruent.

Table 1

Concepts and Principles (con't)

35. Theorem: If one acute angle of a right triangle is 30°, the side opposite this angle is one half the hypotenuse.

36. Synthetic vs. analytic models of proof

37. Theorem: A point equidistant from the end points of a line segment lies on the perpendicular bisector of the line segment.

38. Axial symmetry

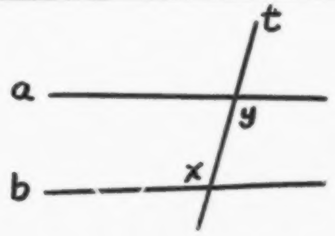
39. Central symmetry

40. Locus

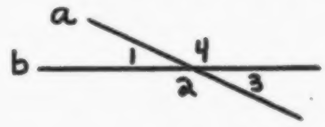
Table 2

Symbol Combinations with Examples

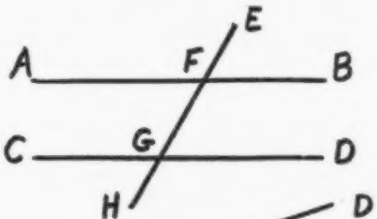
1. A single lower case letter represents a line or an angle.



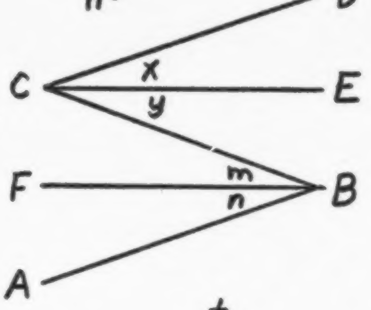
2. A single lower case letter represents a line and a single numeral represents an angle.



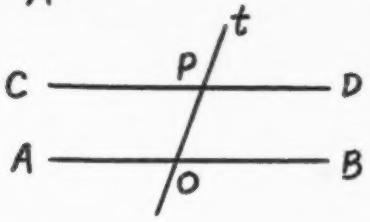
3. Two upper case letters represent lines and three upper case letters represent angles.



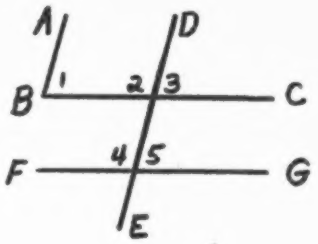
4. Three upper case letters represent some angles and a single lower case letter represents other angles in the same figure.



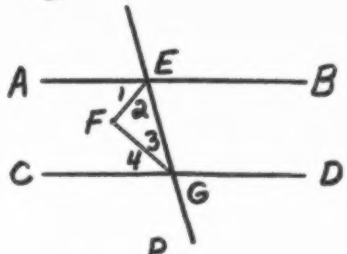
5. Two upper case letters represent some lines and a single lower case letter represents other lines in the same figure.



6. Two upper case letters represent a line and a single numeral represents an angle.



7. Three upper case letters represent some angles and a single numeral represents other angles in the same figure.



8. Three upper case letters represent some angles, a single lower case letter represents some angles, and a single numeral represents other angles all in the same figure.

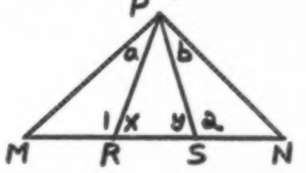


Table 3

Geometric Figural Context

1. Two intersecting lines
2. Two parallel lines
3. Two lines intersecting on the right with a transversal
4. Two lines intersecting on the left with a transversal
5. Two horizontal parallel lines intersected by a transversal
6. Two vertical parallel lines intersected by a transversal
7. Two vertical parallel lines intersected perpendicularly by a transversal.
8. Two horizontal parallel lines intersected by two transversals
9. Two pairs of parallel lines each intersecting the other pair
10. Three parallel lines
11. Two horizontal parallel lines intersected perpendicularly by a transversal
12. A triangle with one exterior angle bisected
13. Two angles with sides parallel respectively, right side to right side and left side to left side
14. Two angles with sides parallel respectively, right side to left side and left side to right side
15. Two similar triangles
16. One right and one obtuse triangle
17. A triangle with sides extended to form six exterior angles
18. A pair of congruent right triangles
19. Two angles with sides perpendicular respectively, right side to right side and left side to left side
20. Two angles with sides perpendicular respectively, right side to left side and left side to right side
21. An isosceles triangle

Table 3

Geometric Figural Context (con't)

22. An equiangular triangle
23. A pair of isosceles triangles
24. A $30^\circ - 60^\circ$ right triangle
25. An isosceles right triangle
26. A pair of congruent obtuse triangles
27. Figures with a single axis of symmetry
28. Figures with more than one axis of symmetry
29. Figures with central symmetry
30. The locus of points equidistant from two parallel lines
31. The locus of points equidistant from the sides of an angle
32. The locus of points equidistant from two given points
33. The locus of points a given distance from a given point

Table 4

Sequence of Ordered Triples (Concept, Symbols, Figure) Included in the Linear and Branch Programs

Ordered Triple	Frame #		Ordered Triple	Frame #		Ordered Triple	Frane #	
	Linear	Branch		Linear	Branch		Linear	Branch
(1,1,2)	7	5	(18,5,5)	215	170	(30,3,22)	589	451
(2,1,2)	13	5	(19,2,5)	218	173	(18,3,22)	598	453
(3,1,3)	16	13	(17,2,5)	222	176	(24,7,21)	619	468
(4,1,4)	20	21	(18,2,5)	230	177	(30,7,21)	621	467
(5,1,4)	20	21	(20,2,5)	231	177	(19,7,21)	635	476
(6,1,4)	23	25	(19,2,5)	237	184	(23,3,21)	642	481
(7,1,4)	25	29	(21,2,11)	243	185	(30,7,21)	655	489
(8,1,4)	28	33	(19,2,11)	247	188	(23,6,23)	662	501
(4,2,5)	31	37	(17,2,5)	266	201	*(32,-,-)	699	541
(5,2,5)	33	37	(7,2,5)	273	201	*(33,-,-)	730	561
(6,2,5)	35	37	(17,1,5)	278	214	(34,3,18)	766	597
(7,2,5)	37	37	(17,2,6)	288		(28,3,18)	771	601
(8,2,5)	38	37	(20,3,9)	299	221	(35,7,24)	773	605
(4,1,6)	39	41	(19,6,9)	316	237	(27,3,24)	778	612
** (5,1,6)	40	41	(19,6,12)	328	241	(35,3,24)	781	615
(6,1,6)	42	41	(17,6,12)	330	245	(26,7,25)	785	624
(7,1,6)	46	41	(22,4,13)	338	261	(23,7,24)	797	637
(8,1,6)	48	41	(15,4,13)	341	261	(36,4,26)	798	641
(4,3,5)	54		(2,4,13)	343	261	(36,3,26)	810	653
(6,3,5)	55	49	(19,4,13)	347	263	(36,8,26)	812	661
(5,3,5)	56	49	(15,4,14)	365	269	(36,8,21)	816	677
(7,3,5)	57		(17,4,14)	372	273	(36,7,21)	825	689
(8,3,5)	58	49	(20,4,14)	376	279	(19,7,21)	827	689
(10,1,5)	60	53	(23,4,8)	383	285	(30,7,21)	830	700
(9,1,5)	65	57	(15,4,8)	384	285	(37,3,21)	833	701
(11,1,5)	68	65	(17,4,8)	394	293	(35,7,24)	840	713
(12,1,5)	83	69	(24,3,15)	399	301	(23,7,24)	844	721
(13,1,5)	96	85	(23,3,15)	402	303	(35,3,24)	862	729
(14,1,5)	109	97	(25,3,16)	413	305	(26,6,21)	865	736
(4,3,8)	127	113	(10,3,16)	414	309	(30,6,21)	868	738
(13,3,8)	129	113	(9,3,16)	417	312	(11,6,21)	875	741
(3,4,9)	132	117	(26,7,17)	451	333	(38,1,27)	877	749
(12,4,9)	134	117	(23,7,17)	453	334	(38,5,28)	879	749
(11,4,9)	140	125	(27,1,6)	460	337	(39, ,29)	882	757
(6,4,9)	144	133	(23,1,6)	464	340	(40,3,30)	890	769
(13,4,9)	151	137	(28,3,18)	469	341	(40,3,31)	892	771
(15,1,1)	161	145	(24,3,18)	472	349	(40,3,32)	893	773
(16,1,10)	163	149	(12,7,12)	525	405	(40,3,33)	897	775
(10,1,10)	165	149	(11,7,12)	527	405	(23,3,21)	941	837
(2,1,10)	165	149	(26,7,12)	533	409	(35,3,24)	942	841
(15,1,10)	173	153	(29,7,19)	542	417	(17,4,8)	943	844
(17,5,5)	180	157	(27,7,19)	551	423	(23,4,8)	945	844
(11,5,5)	188	161	(30,7,21)	563	429	(20,7,5)	949	849
(15,5,5)	195	165	(24,7,21)	573	433	(23,7,5)	951	852
(9,5,5)	198	165	(31,3,22)	584	449			

*The concepts, inverse and contrapositive, were presented without geometric figures.

**Braces represent related application problems involving the concept indicated.

were supplied and then gradually reduced until the desired response could be emitted without help. The vanishing technique was followed in determining the amount of help given to the students in constructing formal proofs. In the proof of the first proposition, the entire proof was written out for the student. In subsequent proofs, the students were given partial proofs and were expected to complete the missing statements and reasons. Then the students were given a complete analysis and were expected to construct the entire synthetic proof. The last step in the vanishing technique occurred on the criterion test where the students were expected to construct two original proofs without any help.

The vanishing technique was also used in the building of concepts by gradually reducing the prompts. Consider, as an example, the concept of "alternate interior angles," formed by a transversal intersecting two given lines. The term is first defined denotatively in frame 23 and an example is required using the same figure.

23

In figure 6, a pair of nonadjacent interior angles on opposite sides of the transversal t are z and s. We call z and s a pair of alternate interior angles. In figure 6, the other pair of alternate interior angles are _______ and _______ (ltr).

Alternate interior angles are next discussed in frame 35. A new figure is used to illustrate the concept, and numerals are used to represent the angles instead of single letters as in the first example. This time both pairs of alternate interior angles are requested.

35

In figure 8 there are two pairs of alternate interior angles. One pair is ∠ ________ (#) and ∠ ________ (#). The other pair is ∠ ________ (#) and ∠ ________ (#).

Alternate interior angles are next discussed in frame 42. Again, a new figure is used and part of the name of the concept is asked for instead of examples only, as was done in the two previous frames.

42

In figure 9, we can say that ∠ r and ∠ w are a pair of ________ (wd) interior angles.

The situation is reversed in frame 43 where the concept is named and an example is requested.

43

In figure 9, the other pair of alternate interior angles is ________ and ________ (ltr).

It is not until frame 55 that both words, "alternate interior", are elicited in the same frame. Again, the concept is illustrated with a new figure in which the angles are named by three capital letters.

55

∠ BFG and ∠ CGF are a pair of ________ ________ (wds) angles formed by the transversal EH intersecting the two lines AB and CD.

The previous series of frames illustrate another technique advocated by Skinner, that of building a network or web of associations. The concepts and principles in the program are introduced in one figural context and then illustrated in other contexts to increase generality.

In presenting the steps of the formal proofs in the program, the logical model of the syllogism was followed extensively. To illustrate:

103

One of the first propositions you proved in plane geometry was that if two angles are supplements of the same angle, they are _______ (wd).

104

Therefore, since $\angle y$ and $\angle z$ are both supplementary to $\angle x$, we can conclude in statement 3 that $\angle y$ and $\angle z$ are _______ (wd).

Consecutive frames are presented vertically here for convenience. The student worked horizontally, not vertically down the page, in both the linear and the branch programs. Syllogisms were used also in discussing a plan of proof to use in a particular situation. To illustrate:

145

Theorem 1 states: "If two lines form equal alternate interior angles with a transversal, _______ _______ _______ _______ (wds)."

146

Therefore, in problem 3, if we can prove that $\angle y$ equals $\angle m$, we can conclude: CE is _______ _______ _______ (wds. and ltrs).

Construction of the Branch Program

The linear program was written first. The branch program was written to cover the same concepts and principles as those presented in the linear program. The difference was in the size of the steps in the programs and the format of the questions. The branch program covered the material with 179 questions. The linear program had 916 questions. The branch program included multiple-choice questions exclusively while the linear program had completion questions only. Three alternatives were presented for each multiple-choice question in the branch program.

Branch Frame

An example of the format of the frames in the branch program is given below.

341

We proved in corollary 4 that the acute angles of a right triangle are complementary. Turn to figure 47 for a statement of corollary 5. This corollary gives us another way to prove that two triangles are congruent. Which statement below is the best description of congruent figures?

a. Congruent figures have the same shape. Frame 342.

b. Congruent figures have the same size. Frame 343.

c. Congruent figures have the same size and shape. Frame 344.

The student read the frame, wrote the letter of the alternative he chose on the answer sheet, and then turned to the frame in the booklet indicated after that alternative. If the alternative chosen was correct, the student was so informed and was then directed to the frame containing the next question in the program. If a student had chosen alternative 'c' to the question above, he would have turned to frame 344 which is reproduced below.

344

Your answer was c: "Congruent figures have the same size and shape." You are right. Both these conditions of identical size and shape must be met before we can say that two figures are congruent. Go on to 345.

If the alternative chosen to a particular question was wrong, the student was given an explanation of the error and directed to go back to the frame containing the question and try again. For example, if a student had chosen alternative 'a' to the question in frame 341, he would have turned to frame 342 which is reproduced below.

342

Your answer was a: "Congruent figures have the same shape." This is true but it is not an accurate description. Two squares have the same shape but one might have sides 2 inches long and the other have sides 4 inches long. They would have the same shape but not be congruent. Go back to 341 and choose again.

The program frequently asked students to write in the reasons for the statements of a proof in the Figures Booklet and then to turn to a particular frame to check their work. If they completed the proof correctly they were sent ahead. If not, an explanation of their errors was provided. This procedure is illustrated below.

737

The last two reasons of the proof in figure 84 are given below. Check whether or not you got them right.

Reasons

4. Subtraction Axiom

5. Theorem 6: If two angles of a
triangle are equal, the sides
opposite these angles are equal.

Write the letter(s) of the statement(s) below which applies to you in the answer column.

a. If you missed reason #4, go to Frame 739
b. If you missed reason #5, go to Frame 738
c. If you got both reasons correct, go to Frame 740

The nature and amount of explanation given to each student was determined by the errors he made in working through the program, which is the procedure Crowder advocates.

The same Figures Booklet was used for both programs so that students in all four treatment groups were given the same illustrative materials. Also, the concepts and principles were covered in the same sequence in the two programs.

Branch Format

A block of four consecutive frames was allotted for each multiple-choice question and its three alternatives. The question was always presented in the first frame of the block. A table of random numbers (9:366-370), was used to determine which of the three alternatives, a, b or c would contain the correct answer. In addition, a table of random numbers was used to determine whether the correct answer would be contained in the 2nd, 3rd or 4th frame in each four-frame block.

Branch Booklets

The seven booklets containing the branch program were 8-1/2" by 11" in size and each had 60 pages except the last booklet which had 66 pages. Each page contained two frames, one on the top half of the page and one on the bottom half. The student worked through the 60 frames on the top half of each booklet and then started on the bottom frames.

The Criterion Achievement Test

It is desirable to use a standardized criterion test in measuring achievement in a school subject area since validity and reliability data are available as well as norms for comparison purposes. A search of the test literature revealed no standardized test covering such a small segment of plane geometry as was included in this unit on parallel and perpendicular lines. The linear and branching programs covered exactly the same concepts as were presented in the chapter entitled "Parallel and Perpendicular Lines" in the text, New Plane Geometry by Welchons and Krickenberger (24). A series of achievement tests had been prepared by the authors explicitly related to the material presented in their text. Therefore, the decision was made to use items from the published test covering the material on parallel and perpendicular lines as the criterion test for this experiment.[6] The 54 test items chosen included the following types: true-false, completion, applications, supplying reasons and proofs. Some of the items included two parts, so each item was given a value of two points to avoid fractional scores. Therefore, the range of possible points on the criterion achievement test was 0 to 108.

---

[6] Written permission was requested and obtained from the publisher, Ginn and Co., to program a portion of the text, New Plane Geometry, and to use items from the related achievement test in this study.

The criterion test sampled 27 of the 40 concepts and principles presented in the programed unit on parallel and perpendicular lines. A measure of reliability was obtained by the split-half, odd-even method corrected for length by the Spearman Brown Formula:

$$r_c = \frac{2r_{1,2}}{1 - r_{1,2}}$$

Computed in this manner, the Pearson product moment correlation coefficient was .92. The subjects used to obtain these reliability data were a class of students in plane geometry who had finished the chapter on parallel and perpendicular lines in the same textbook as the one programed in the experiment. This class was in the same high school as those participating in the study, but the class had a different teacher and was not included in the experiment. A Pearson product moment correlation coefficient was also computed in the same manner using the scores of the four experimental groups on the criterion test given at the end of the experiment. With this group, $r = .93$.

The items on the criterion test were divided into four types and an item analysis performed using the Davis Item Analysis Chart. (8) The four categories into which the items were placed are:

1. retention: items that appeared in the programs
2. numerical applications: items not in the programs and involving numerical computations
3. verbal application: items not included in the programs but involving the same concepts and not requiring numerical computations

4. verbal reasoning and proofs: items in which the student is required to tell why a relation holds. This includes items which are steps of formal proofs, where the proofs are not included in the programs.

The verbal reasoning and proof items were more difficult for the students than the other types. Pure recall and numerical application items were the easiest. The table also indicates that the more difficult the item, the better the discrimination. The data for this item analysis were taken from the scores on the post-test of the subjects using programed materials.

Item Analysis Table

Category of Item	N	Mean Difficulty Index*	Mean Item Discrimination Index**
retention	7	72	31
numerical application	9	71	33
verbal application	19	61	32
verbal reasoning and proofs	19	58	37

* The range of the difficultyindex is from 1 to 100. The lower the index the more difficult the item.

** The range of the discrimination index is from 1 to 100. The higher the index the better the discrimination.

The Attitude Questionnaire

The attitude questionnaire was composed of 13 items and was divided into three sections. The first seven items constituted the first section. Each item in the first section contained five alternative responses from which the student was instructed to choose one. The alternatives were scaled for each item from a very unfavorable attitude toward programed instruction to a very favorable attitude toward programed instruction. In scoring this section, the five alternatives for each item were weighted 5, 4, 3, 2 or 1, with a weight of 5 as the most favorable attitude toward programed instruction. The alternative least favorable to programed instruction was given a weight of 1. A score on attitude toward programed instruction was obtained for each subject in the study by summing the weights of the alternatives he chose on the first seven items of the attitude questionnaire. The possible range of scores was from 7 to 35. The higher the score the more favorable was the attitude toward programed instruction.

The second section of the attitude questionnaire was composed of questions 8, 9 and 10. These items were included to obtain a comparison between the linear and the branch programs. Weights were assigned to the alternatives so that a high score on this section of the questionnaire would indicate a high preference for the branch program and a low score would indicate a preference for the linear program. An alternative which indicated no preference for either of the programs was assigned a weight midway between the two extremes. In item 8, the alternatives were assigned the following weights: a - 4, b - 0, c - 2. In item 9, the alternatives were assigned the following weights: a - 5, b - 1, c - 3, d - 3, e - 4, f - 2, g - 3. Alternatives "a" to "g" exhausted the possible combinations. However, "h" was included for any subject who might misinterpret the other alternatives and feel that the combination he wanted was not included. In item 9, consideration was given also to the number of the choice of each alternative checked. The weight for a particular alternative was multiplied by three if it was the first choice, two if it was the second choice and one if it was the third choice.

A preference score for the linear vs. the branch program was obtained by adding the weight of the alternative chosen in item 8 to the sum of the products of the weights times the choice values for the alternatives checked in item 9. Item 10 was not included in the score. The range of the preference score was 10 to 30. The alternatives for items 8 and 9 are given below, listed in order from high preference for the branch program to high preference for the linear program. The weights are indicated.

8. Next week in studying plane geometry, I want to:

    a. 4 use the multiple-choice program
    c. 2 go back to the usual procedure of having a teacher and regular textbook
    b. 0 use the completion question program

9. For the rest of this semester in plane geometry I want to: (Indicate 3 choices: 1st choice: 1; 2nd choice: 2; 3rd choice: 3).

    a. 5 use the multiple-choice program all the time
    e. 4 have a teacher and regular textbook part of the time and use a multiple-choice program on my own part of the time.
    c. 3 be in a class with a teacher and use a regular textbook.
    d. 3 use programed materials but alternate between the multiple-choice and completion programs.
    g. 3 alternate all three types of instruction: 1) teacher and textbook, 2) multiple-choice program and 3) completion program
    f. 2 have a teacher and regular textbook part of the time and use a completion question program part of the time
    b. 1 use the completion type program all the time

Alternative"c" in item 8 and "c," "d" and "g" in item 9 indicate no preference for either the linear or the branching program. Therefore, they were assigned weights "2 and 3" respectively, values midway between the two extremes for each item.

An example of the most extreme preference for the branching program would be scored as follows: Such a subject would check "a" in item 8; this alternative would be scored as 4 points. In item 9, he would select alternatives "a" as his first choice, "e" as his second choice, and "c", "d" or "g" as his third choice. His score would be 5 x 3 = 15 for his first choice, 4 x 2 = 8 for his second choice, and 3 x 1 = 3 for his third choice. Adding these four subscores together would give him a total preference score of 30. A subject who chose only alternatives with no preference for either the linear or the branching program would receive a preference score of 20. A subject who wanted to indicate an extreme preference for the linear program could receive a score as low as 10. The attitude score and the preference score were the only quantitative measures obtained from the attitude questionnaire.

One of the main criticisms of attitude questionnaires given at the end of a research experiment is their inability to counteract the "Hawthorne Effect". Students enjoy the experience of an experimental situation and when questioned, will generally state a preference for the experimental treatment as opposed to the regular routine. In the construction and administration of the attitude questionnaire in this experiment, a concerted effort was made to overcome this experimental bias. Students were deliberately not told how long the experiment would last. When given the attitude questionnaire, the subjects were told that their choices would have some bearing on the type of programed materials, if any, that they would receive the following week. Therefore, if a student did not like the program on which he was working, there was no reason for indicating that he preferred it, since he knew this might result in his being given the same program the following week, when another choice was available.

Pilot Study of the Programed Materials

The two programs were pretested by six volunteer students enrolled in plane geometry at Urbana High School, Urbana, Illinois. Three students worked on each program. The students indicated faulty or ambiguous frames as they worked through the programs. The consensus of those using the linear program was that the material moved too slowly. On the basis of the pilot study and the written comments of otherswho examined portions of the programs, each program was completely revised before final typing for the main study at Rantoul High School.

CHAPTER IV.

RESULTS

The four experimental treatment groups were compared on the variables of general mental ability, achievement in algebra, and achievement in geometry. Deviation IQ scores were obtained from the Henmon-Nelson Test of Mental Ability administered the week before the experiment began. The final freshman algebra course grades in the form of percentiles were obtained from the permanent records maintained in the school office. The first quarter of the school year ended the same week the experiment started, so first quarter geometry grades for all students in the study were available directly from the geometry teacher.

The criterion achievement test was administered before the experiment began, but the scores are not included in the analysis due to their unreliability. Five individuals were dropped from the experiment because they moved from the school district before all the data were collected. One person was transferred from Group 2 to Group 3 so that each subgroup would contain the same number of students.

The means and standard deviations for all treatment groups and subgroups on the three independent variables are presented in Table 5.

Analysis of Data on the Independent Variables

A two way analysis of variance was used to test the hypothesis that no differences existed between the means of the four groups on the three independent variables. Four experimental treatments and

Table 5

Mean Scores of Treatment Groups and Ability Subgroups On The Variables of IQ Scores, Algebra Grades and Geometry Grades

A. Deviation IQ Scores

Group	Total Group			High Ability Subgroup			Low Ability Subgroup		
	N	Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.
1	12	110.8	11.1	6	119.5	4.2	6	102.2	8.6
2	12	116.9	11.3	6	125.0	9.8	6	108.8	5.3
3	12	116.0	11.8	6	123.5	12.2	6	108.5	4.8
4	12	113.0	11.9	6	121.5	8.7	6	104.5	8.0

B. Final Freshman Algebra Course Grades

Group	Total Group			High Ability Subgroup			Low Ability Subgroup		
	N	Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.
1	12	83.6	9.2	6	89.2	5.2	6	78.0	9.2
2	12	87.5	7.2	6	89.5	3.9	6	85.5	9.5
3	12	89.0	5.5	6	88.0	7.1	6	90.0	3.9
4	12	84.3	9.7	6	89.3	6.1	6	79.3	10.5

C. First Quarter Plane Geometry Course Grades

Group	Total Group			High Ability Subgroup			Low Ability Subgroup		
	N	Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.
1	12	87.3	8.1	6	93.0	4.9	6	81.5	6.4
2	12	89.6	5.2	6	91.3	5.3	6	87.8	4.9
3	12	89.7	5.8	6	91.5	4.8	6	87.8	6.6
4	12	87.1	8.7	6	90.0	8.2	6	84.2	9.0

two ability levels were compared. The raw scores were punched into IBM cards, and the IBM 1401 computer system at the University of Illinois Statistical Services Laboratory was used to obtain the analysis of variance data.

Certain assumptions must be made in applying the two way analysis of variance to interpret the results of an empirical study. 1) Each treatment group and subgroup originally was a representative sample from a specified population. This means that each treatment subgroup was originally drawn at random from the corresponding level of the given population, and the number drawn at each level was proportional to the number of individuals at that level in the population. This assumption was met by ordering the scores on the control variable, general intelligence, for all students in the two classes from which the treatment samples were drawn. Then individuals within each level on the control variable were randomly assigned to the treatment groups by means of a table of random numbers. The exact procedure was described in Chapter III. 2) The distribution of criterion measures for the subpopulation corresponding to each treatment subgroup is a normal distribution. The Norton study, reviewed in Lindquist (14:78-86), presents empirical evidence that even marked departure from a normal distribution will have little effect on the F-test used in analysis of variance. 3) Each of the subgroups are samples from populations with the same variance. This is commonly referred to as the "homogeneity of variance" requirement and can be tested by comparing the sample variances. The Hartley F (max) Test, described in Walker and Lev's Statistical Inference (23:193), was used to test the assumption of homogeneity of variance on each of the criterion measures. No significant differences were found, so the assumption was valid in this experiment.

The analysis of variance summaries for the three independent variables are presented in Table 6.

No significant differences existed between the means of the four experimental treatments on any of the independent variables. Therefore, it is reasonable to assume that the four groups began the experiment with approximately the same general ability and background in algebra and geometry.

Significant differences did exist between the high and low ability subgroups on all three independent variables. The differences were significant at the .01 level. We can reasonably assume that the high ability students began the experiment with greater general ability as measured by the Henmon-Nelson Mental Ability Test and greater achievement in algebra and geometry as measured by course grades than the low ability students. Table 6 also reveals that no significant interaction existed between ability level and the program treatment to which the students were assigned on any of the three independent variables.

Analysis of Data on Achievement

This experiment was designed to measure the effects of programed instruction in plane geometry on the variables of achievement, retention and attitude toward programed instruction. The effectiveness of the treatments was compared on the basis of mean scores of the groups on each of the dependent variables. The results of the four experimental groups on the criterion achievement test administered at the end of the

Table 6

Analysis of Variance Summary for the Independent Variables

A. Deviation IQ Scores

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	6139.31	47		
Between Treatments	280.73	3	93.58	1.42
Between Ability Levels	3217.69	1	3217.69	48.92**
Interaction	9.73	3	3.24	.05
Within Subgroups	2631.17	40	65.78	

B. Final Freshman Algebra Course Grades

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	3118.48	47		
Between Treatments	237.90	3	79.30	1.48
Between Ability Levels	402.52	1	402.52	7.50**
Interaction	331.56	3	110.52	2.06
Within Subgroups	2146.50	40	53.66	

C. First Quarter Plane Geometry Course Grades

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	2303.48	47		
Between Treatments	72.73	3	24.24	.59
Between Ability Levels	450.19	1	450.19	10.88**
Interaction	125.73	3	41.91	1.01
Within Subgroups	1654.84	40	41.37	

**Significant. (probability less than .01)

of the experiment are given in Table 7.

Table 7

Mean Scores on Achievement Post-Test

Group	Total Group		High Ability Subgroup		Low Ability Subgroup	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
1	74.3	17.8	83.7	14.6	64.8	16.5
2	72.6	17.5	80.7	13.2	64.5	18.6
3	77.2	19.5	85.8	15.2	68.5	20.7
4	72.0	18.5	81.8	15.1	62.2	17.2

A two way analysis of variance was used to test Hypothesis 1, viz., that no significant differences would exist between the means of the four experimental groups on achievement. Hypothesis 2, viz., that no differences in achievement would exist between the two ability levels, was tested by the same analysis of variance. Each of the stated hypotheses in the study was tested with $\alpha = .05$. The summary of the analysis is contained in Table 8.

Table 8

Analysis of Variance of Achievement Scores

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	15036.00	47		
Between Treatments	193.17	3	64.39	.24
Between Ability Levels	3888.00	1	3888.00	14.22**
Interaction	21.83	3	7.28	.03
Within Subgroups	10933.00	40	273.33	

**Significant. (probability less than .01)

Based on the analysis, the hypothesis of no differences in mean achievement scores for the four experimental treatments was accepted. The hypothesis that no differences would exist in achievement between the high and low ability subgroups was rejected. As measured by the criterion achievement test, the high ability students learned significantly more (.01 level) than the low ability students.

Analysis of Data on Retention

The same criterion achievement test was administered to all subjects in the study seven weeks after the experiment was completed as a measure of retention. Five weeks of this intervening time was spent in regular classroom instruction covering chapters in the text on polygons and circles. The Christmas vacation accounted for the other two weeks of the intervening period.

The mean scores and standard deviations on the delayed achievement test are given in Table 9.

Table 9

Mean Scores on Delayed Achievement Test

Group	Total Group		High Ability Subgroup		Low Ability Subgroup	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
1	72.1	21.6	83.2	18.4	61.0	19.9
2	78.3	17.6	85.8	12.1	70.8	20.0
3	80.3	15.1	87.0	9.5	73.5	17.3
4	69.8	20.8	80.2	20.0	59.5	17.2

A two way analysis of variance was used to test the existence of differences between the means of the four experimental groups on the delayed achievement test as well as differences between mean scores of high and low ability subgroups. This is a test of hypotheses 3 and 4, respecitvely. The summary of the analysis is contained in Table 10.

Table 10

Analysis of Variance of Delayed Achievement Scores

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	16703.25	47		
Between Treatments	885.75	3	295.25	1.00
Between Levels	3816.33	1	3816.33	12.89**
Interaction	160.83	3	53.61	.18
Within Subgroups	11840.34	40	296.01	

**Significant. (probability less than .01)

The analysis of variance revealed no significant differences in mean delayed achievement scores for the four experimental groups and no significant interaction. Therefore, the null hypothesis of no difference in retention between the treatment groups was accepted.

Significant differences did exist between the high and low ability levels. Table 9 reveals that under each treatment, the high ability subgroup did considerably better on the delayed achievement test than the low ability subgroup; therefore, Hypothesis 4, viz., that no differences in retention would exist between the ability levels, was rejected.

The use of raw scores on a delayed achievement test as a measure of retention has certain weaknesses. One confounding factor is that students are free to discuss test questions during the intervening period. Also, the material covered during the intervening period of time may give added practice in applying the concepts covered in the experiment. Therefore, a retention score was also computed, considering only those items which were correct on the post-test. The score is the proportion of items correct on the post-test which were also correct on the delayed test given seven weeks later. This retention score is a measure of how much the student remembered of what he knew at the end of the experiment. The main assumption underlying this score is that the items a student answered correctly on either administration of the test were the items to which he knew the answer. The range of this retention score is from 0 to 100. This measure will be referred to as the retention score to distinguish it from the delayed achievement score discussed earlier. The mean retention scores for all treatment groups and subgroups are given in Table 11. Both mixed-treatment groups had better mean retention scores than either group which used one program exclusively.

Table 11

Mean Retention Scores

Group	Total Group		High Ability Subgroup		Low Ability Subgroup	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
1	76.9	15.2	82.7	14.9	71.2	14.5
2	86.1	11.0	89.3	6.2	82.8	14.2
3	86.2	8.9	89.2	5.2	83.2	11.2
4	79.8	13.9	85.3	10.5	74.3	15.5

A two way analysis of variance was used to test Hypotheses 3 and 4, viz., that no significant differences would exist between the mean retention scores of the four treatment groups or between the two ability levels. The summary of the analysis is presented in Table 12.

Table 12

Analysis of Variance of Retention Scores

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	7645.00	47		
Between Treatments	771.83	3	257.28	1.75
Between Ability Levels	918.75	1	918.75	6.25*
Interaction	75.75	3	25.25	.17
Within Subgroups	5878.67	40	146.97	

*Significant. (probability less than .05)

Hypothesis 3, viz., that no differences would exist in mean retention scores between the four experimental treatments, was accepted. However, a significant difference at the .05 level did exist between the high and low ability levels.

An examination of Table 11 reveals that the high ability subgroup did considerably better than the low ability subgroup under each of the four treatments. Therefore, the hypothesis of no difference in retention between the ability levels was rejected.

Both measures of retention produced similar results in this experiment. The raw scores on the delayed achievement test and the retention scores, indicating the proportion of items correct on the posttest which were also correct seven weeks later, both indicated significant differences between ability levels but not between treatments.

Summary of Data on Attitude Toward Programed Instruction

The attitude questionnaire was administered three times during the study - halfway through the program, at the end of the program and seven weeks later. The range of possible scores is from 7 to 35. A high attitude score indicates the student has a favorable impression of programed instruction and prefers it to classroom instruction with a teacher and regular text. A low score indicates the student prefers the regular classroom instruction to programed instruction.

The mean attitude scores for the four treatment groups and the ability subgroups for each administration of the questionnaire is given in Table 13.

All groups had mean attitude scores favorable to programed instruction halfway through the experiment and at the end of the experiment. A score of 21 is interpreted as indicating no preference for either method of instruction. At the end of the experiment, time 2, the high ability subgroups expressed attitudes more favorable to programed

Table 13

Mean Scores of Attitude Toward Programed Instruction

A. Time 1: (Halfway Through Experiment)

Group	Total Group		High Ability Subgroup		Low Ability Subgroup	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
1	28.3	4.5	30.3	3.3	26.2	4.8
2	24.9	7.7	24.8	7.0	25.0	6.1
3	24.3	6.1	26.3	6.7	22.3	5.1
4	25.8	6.8	24.7	6.5	26.8	7.4

B. Time 2: (End of Experiment)

Group	Total Group		High Ability Subgroup		Low Ability Subgroup	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
1	26.0	5.8	27.8	5.0	24.2	6.5
2	25.1	6.8	26.8	4.8	23.3	8.4
3	22.8	7.0	24.2	8.0	21.3	6.3
4	25.1	6.6	25.2	6.1	25.0	7.5

C. Time 3: (Seven Weeks After Experiment)

Group	Total Group		High Ability Subgroup		Low Ability Subgroup	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
1	21.1	5.1	22.3	5.7	19.8	4.7
2	20.9	5.9	20.3	5.5	21.5	6.6
3	19.4	5.4	17.7	2.2	21.2	7.2
4	22.2	7.6	22.0	6.8	22.3	9.1

instruction than the low ability subgroups under all four treatments. Seven weeks later all treatment groups had mean attitude scores less favorable to programed instruction than at the end of the experiment. The mean attitude scores for the four experimental groups at each administration of the questionnaire are pictured graphically in Figure 1.

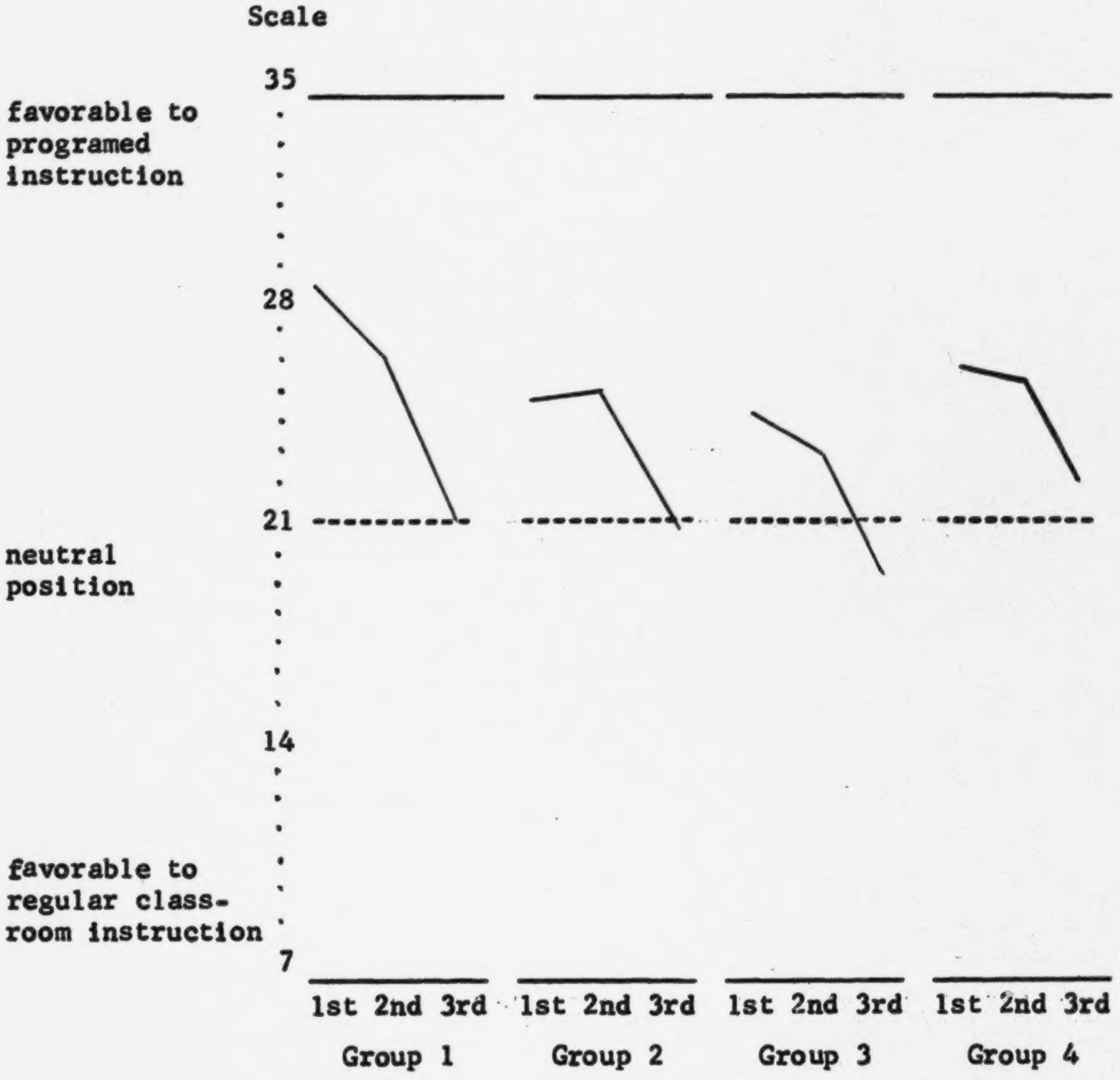
In Figure 1 the trend for each of the four groups is toward the neutral position from an earlier rather favorable attitude toward programed instruction. The time interval between the first and second administrations of the attitude questionnaire was one week while the time interval between the second and third administrations was seven weeks.

One pertinent observation from Figure 1 is that at no time did the students reject programed instruction by expressing a highly favorable attitude toward regular classroom instruction. This is true even after seven weeks back in the regular classroom situation. The initial high preference for programed instruction disappeared after a period of time but the pendulum of attitude did not swing to the other extreme.

Figure 1

Mean Attitude Scores Comparing Programed Instruction with Traditional Classroom Instruction

(Range 7-35)



Analysis of Attitude Data

A two way analysis of variance was used to test for significant differences in attitudes toward programed instruction between the four experimental groups and the two ability levels. No significant differences in mean attitude scores were revealed between experimental treatments or ability levels at any of the three times that the attitude questionnaire was administered. Also, there was no significant interaction between treatment or ability. This information is contained in Table 14. The distributions of responses to individual items in this section of the attitude questionnaire are presented in APPENDIX E.

Summary of Data Indicating A Preference for the Linear or Branch Program

The next analysis was a comparison of attitudes toward the linear versus the branch program. The comparison was made at the end of the experiment. To avoid confusion between this measure and the attitude score previously discussed, the score indicating a preference for the linear or branch program will be referred to as the "preference score". The range of possible scores on this section of the attitude questionnaire is from 10 to 30. A high score indicates a strong preference for the branch program and a low score indicates a strong preference for the linear program.

The mean preference scores for the two mixed-treatment groups and their ability subgroups are given in Table 15. Groups 2 and 3 are the only groups who worked with both programs and therefore, were in a position to make a valid comparison of the two programs.

Table 14

Analysis of Variance of Attitude Scores

A. First Administration of Attitude Questionnaire

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	1893.31	47		
Between Treatments	107.23	3	35.74	.86
Between Ability Levels	25.52	1	25.52	.61
Interaction	88.73	3	29.58	.71
Within Subgroups	1671.84	40	41.80	

B. Second Administration of Attitude Questionnaire

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	2019.48	47		
Between Treatments	69.40	3	23.13	.50
Between Ability Levels	77.52	1	77.52	1.68
Interaction	22.73	3	7.58	.16
Within Subgroups	1849.84	40	46.23	

C. Third Administration of Attitude Questionnaire

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	1676.48	47		
Between Treatments	46.06	3	15.35	.39
Between Ability Levels	4.69	1	4.69	.12
Interaction	55.23	3	18.41	.47
Within Subgroups	1570.50	40	39.26	

Table 15 reveals that both mixed-treatment groups preferred the linear rather than the branch program. A score of 20 was interpreted as indicating no preference for either program. Both high ability subgroups expressed a strong preference for the linear program, the opposite of the preference predicted. The mean preference scores for the two mixed-treatment groups are pictured graphically in Figure 2.

Table 15

Mean Preference Scores

Comparing the Linear and Branch Programs

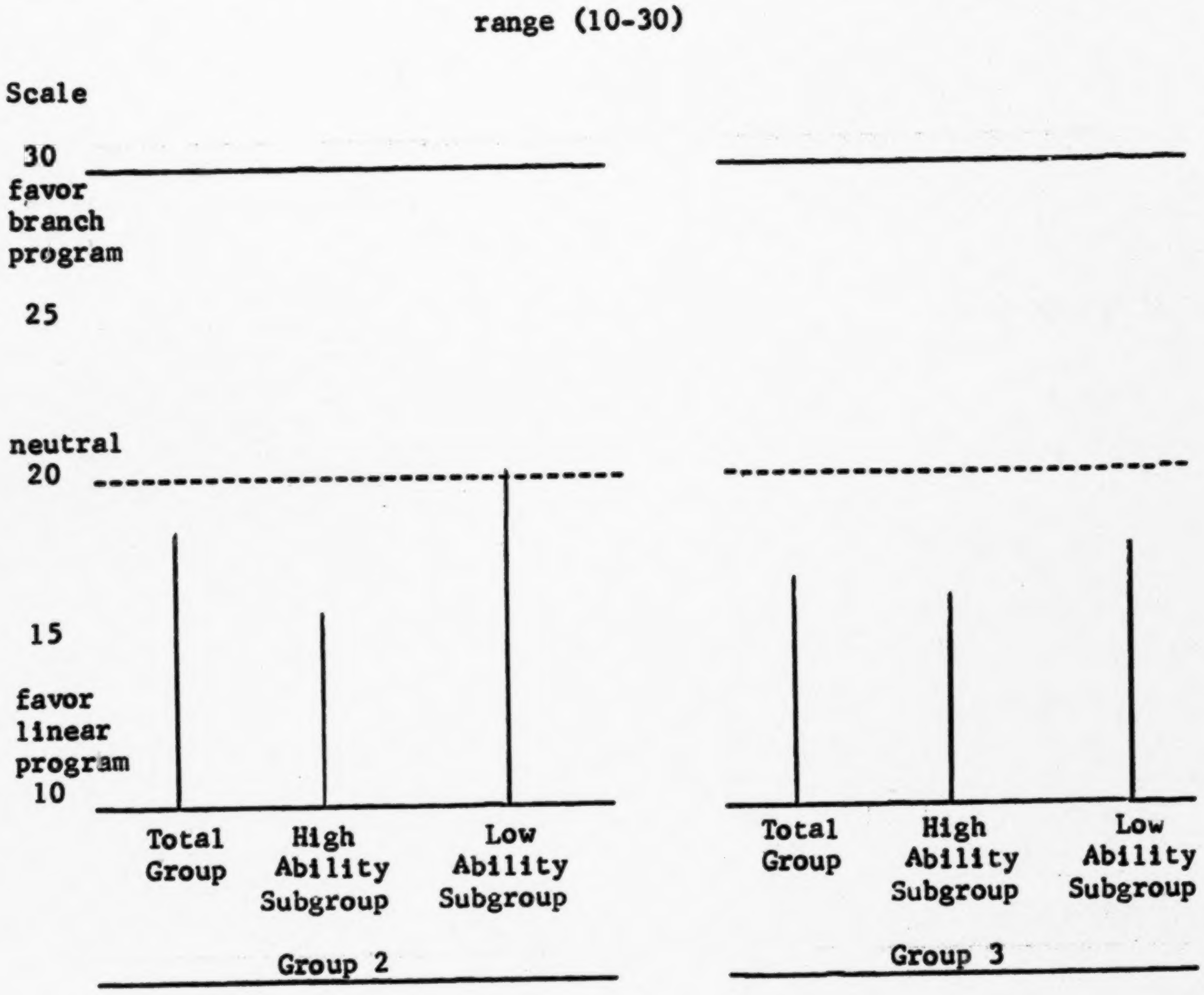
Group	Total Group		High Ability Subgroup		Low Ability Subgroup	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
2	17.8	4.8	15.3	4.8	20.3	3.5
3	16.8	5.0	16.2	5.5	17.3	4.8

A *t*-test (23:147) was used to test whether the means of the mixed-treatment groups or their subgroups were different from the no-preference score of 20. The two mixed groups, which spent approximately the same amount of time on each program, were combined to obtain a larger sample. The hypothesis that the mean of the population from which the sample was drawn is 20 was tested with a two-tailed test at the .05 level. The *t*-test was run for the mixed groups combined and for the high and low ability subgroups of Groups 2 and 3 combined. The results are contained in Table 16.

The mean preference score of the combined mixed-treatment groups was significantly different from the neutral score of 20 in the direction

Figure 2

Mean Preference Scores for the Linear or Branch Program at the End of the Experiment



of preference for the linear program. When the combined high ability subgroups were considered, the difference from the neutral score was also significant in the direction of preference for the linear program. The combined low ability subgroups, those working with both programs, indicated a preference for the linear program, but the difference from 20 was not significant. Therefore, Hypothesis 5, viz., that the high ability students would show no preference for either program, was rejected. Hypothesis 6, viz., that the low ability students would show no preference for either program, was accepted.

Table 16

Significance of Differences from Hypothesized Mean of 20 for Preference Scores of Combined Mixed-Treatment Groups and Ability Subgroups

Sample	N	Mean	S.D.	t	p
Combined Total Groups 2 and 3	24	17.3	4.8	-2.77	<.05
Combined High Ability Subgroups	12	15.8	4.9	-3.00	<.05
Combined Low Ability Subgroups	12	18.8	4.3	- .98	NS

Analysis of Program Errors

The errors on each question in both programs were tabulated and the error distributions are presented in Table 17.

Less than three errors were made on 74% of the questions in the linear program. This is in accord with Skinner's position of keeping the steps small enough that very few errors are made on individual

Table 17

Distribution of Errors for Individual Questions in the Linear and Branch Programs

Linear Program			Branch Program*		
No. of Errors	No. of Questions	% of Questions	No. of Errors	No. of Questions	% of Questions
0	318	35	0	8	4
1	228	25	1	18	10
2	129	14	2	22	12
3	77	8	3	23	13
4	56	6	4	16	9
5	38	4	5	16	9
6	22	2	6	10	6
7	18	2	7	19	11
8	17	2	8	15	8
9-15	13	2	9	3	2
	916	100%	10	9	5
			11	10	6
			12	6	3
			13-19	4	2
				179	100%

*Only first choice errors were considered in arriving at this distribution.

questions. In contrast, less than three errors were made on only 26% of the multiple-choice questions in the branch program, which is in keeping with the point of view that large steps with more errors can contribute to learning if an explanation of each error is included in the program. Very few questions in either program were missed by as many as ten persons which represents approximately 40% of those working on each program. The distribution of errors reveals that most of the questions in both programs had error rates within the range advocated by the two points of view compared in this study.

An examination of individual alternatives to the multiple-choice questions revealed that 13% of the alternatives were not chosen by anyone and therefore, should be eliminated or replaced by other more plausible alternatives. A total of 65% of the alternatives in the branch program were chosen by at least three persons.

The error rates for individual students on both programs were computed. The branch program was intended to be more difficult; therefore, a higher error rate was expected on this program than on the linear program. The error rate for each individual was determined by dividing the total number of errors made by the total number of questions answered. These error rates were averaged to obtain the mean error rate for the group. The results are presented in Table 18.

The evidence is clear that fewer errors were made on the linear than on the branch program by each group. The error rate on the linear program for all groups combined was 7.5. The error rate on the branch program for all groups combined was 20.0. The branch program was nearly three times as difficult as the linear program when error rate is used as a criterion. In general, the high ability subgroup under each

Table 18

Mean Error Rates (in per cent) of Treatment Groups and Ability Subgroups on the Linear and Branch Programs

Linear Program

Group	Total Group		High Ability Subgroup		Low Ability Subgroup	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
2	5.8	3.5	5.3	2.1	6.2	4.7
3	7.9	6.6	5.3	2.9	10.5	8.4
4	8.2	3.8	8.3	4.8	8.0	2.8

Branch Program

Group	Total Group		High Ability Subgroup		Low Ability Subgroup	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
1	21.5	11.4	14.8	9.5	28.3	9.6
2	17.6	8.5	17.2	10.0	18.0	7.7
3	20.3	9.9	18.2	7.7	22.5	12.1

treatment made fewer errors than the low ability subgroup. The one exception was Group 4 on the linear program in which the error rate was slightly less for the low ability subgroup than the high ability subgroup. The difference between the mean error rates of the high and low ability subgroups in Group 1, using the branch program, was significant at the .05 level.

Time and Efficiency of the Experimental Treatments

Table 19

Mean Time (in minutes) Spent on the Programed Materials

Group	Total Group		High Ability Subgroup		Low Ability Subgroup	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
1	442.3	54.2	436.3	71.3	448.2	36.1
2	439.6	62.7	439.0	84.0	440.2	39.8
3	479.4	36.7	475.8	43.3	483.0	32.5
4	511.5	60.5	510.2	49.9	512.8	74.6

The time factor was considered also in judging the value of the experimental teaching programs. The time spent by each student on the programed materials was recorded, and the mean time scores for the four experimental treatment groups and ability subgroups were computed. This information is presented in Table 19.

The mean time spent on the programed materials by Groups 1 and 2 was considerably less than that used by Group 4 which had the linear program exclusively. Very little difference in mean time is indicated between the high and low ability subgroups for any of the treatments.

Again, a two way analysis of variance was used to test the significance of differences in means between treatment groups and between ability subgroups. The summary is presented in Table 20.

Table 20

Analysis of Variance of Times Spent on the Programed Materials

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	172649.3	47		
Between Treatments	41916.2	3	13972.1	4.29*
Between Ability Levels	391.0	1	391.0	.12
Interaction	208.6	3	69.5	.02
Within Subgroups	130133.5	40	3253.3	

*Significant. (probability less than .05)

The analysis of variance revealed that the means of the treatment groups were significantly different at the .05 level while the means of the ability subgroups were not significantly different. No significant interaction was observed between treatment and ability.

A t-test (23:156) was used to test for significant differences in means between pairs of treatment groups. An observed t of 2.96 indicated that the difference in means between Group 1 and Group 4 was significant at the .01 level. Group 2 and Group 4 also had significantly different means at the .01 level with an observed t of 2.86. All other comparisons of mean times between pairs of groups were insignificant at the .05 level. We see that Group 1 using the branch program exclusively covered the material significantly faster than Group 4 using the linear program exclusively.

A measure of efficiency of the various treatments was obtained by forming the ratio of the achievement score to time spent on the programed materials. The assumption is that the higher the ratio, the more efficient is the treatment. These ratios were obtained for each individual, and means were computed for the treatment groups and ability subgroups. The results are presented in Table 21.

Table 21

Mean Efficiency Scores of the
Treatment Groups and Ability Subgroups

Group	Total Group		High Ability Subgroup		Low Ability Subgroup	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
1	16.9	4.4	19.3	3.7	14.5	3.7
2	17.3	5.9	19.3	5.8	15.2	5.7
3	16.0	3.5	17.8	2.3	14.2	3.8
4	14.1	3.1	16.0	2.4	12.2	2.5

Groups 1 and 2 were the most efficient whereas Group 4 was the least efficient. In each treatment group the high ability subgroup was more efficient than the low ability subgroup.

The summary of the analysis of variance to determine the significance of the differences in means of the treatment groups and ability subgroups is presented in Table 22.

The analysis of variance indicated that no significant differences existed between the means of the four treatment groups. However, the means of the high and low ability subgroups were significantly different. No significant interaction between treatment and ability was observed.

Table 22

Analysis of Variance of Efficiency Scores

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	908.8	47		
Between Treatments	72.7	3	24.2	1.54
Between Ability Levels	204.2	1	204.2	12.98**
Interaction	2.4	3	.8	.05
Within Subgroups	629.5	40	15.7	

**Significant. (probability less than .01)

Control Group Data

A third plane geometry class at Rantoul Township High School served as a control group. This class was taught by the instructor who normally taught the two experimental classes. The control group contained 17 students, which made it comparable in size to the experimental groups.

The control group covered the same unit of material on parallel and perpendicular lines, using the regular text, New Plane Geometry. IQ Scores, final freshman algebra grades and first quarter geometry grades were obtained for individuals in the control group in the same manner as for students in the experimental groups. The means and standard deviations of the control and experimental groups on each of the three independent variables are presented in Table 23.

The hypothesis of no differences between the means of the five treatment groups on each of the independent variables was tested by analysis of variance. The results are summarized in Table 24.

The hypothesis of no differences between the five treatment groups on the three independent variables was accepted. The conclusion was made that the five groups were comparable on IQ and achievement in algebra and geometry. Due to the fact that prior to the experiment the students in all five groups had the same geometry teacher, the assumption was made that all groups had received similar instruction in geometry.

A comparison was made also between the control group and the experimental groups on the dependent variables of achievement, retention, amount of time spent on the unit and efficiency. The criterion achievement test was administered to the control group on the same days as it was given to the experimental groups. The control group received ten days of classroom instruction on the unit, which is the same amount of time given to the experimental groups. However, the students in the control group were given homework assignments, while the experimental groups were required to do all their work in class. The time sheets, kept by each student in the control group, revealed that the reported average time spent on homework assignments during the experiment was approximately three-fourths of an hour for every hour of classroom instruction.

Table 23

Means of Control Group and Experimental Groups on the Independent Variables

A. Deviation IQ Scores

Statistic	Experimental Group 1	Experimental Group 2	Experimental Group 3	Experimental Group 4	Control Group 5
N	12	12	12	12	17
Mean	110.8	116.9	116.0	113.0	109.8
S.D.	11.1	11.3	11.8	11.9	10.0

B. Final Freshman Algebra Grades

Statistic	Experimental Group 1	Experimental Group 2	Experimental Group 3	Experimental Group 4	Control Group 5
Mean	83.6	87.5	89.0	84.3	86.6
S.D.	9.2	7.2	5.5	9.7	8.6

C. First Quarter Geometry Grades

Statistic	Experimental Group 1	Experimental Group 2	Experimental Group 3	Experimental Group 4	Control Group 5
Mean	87.3	89.6	89.7	87.1	88.8
S.D.	8.1	5.2	5.8	8.7	7.2

Table 24

Analysis of Variance Comparing the Control and Experimental Groups on the Independent Variables

A. Deviation IQ Scores

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	7966.86	64		
Between Means	519.81	4	129.95	1.05
Within Groups	7447.06	60	124.12	

B. Final Freshman Algebra Course Grades

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	4316.06	64		
Between Means	241.60	4	60.40	.89
Within Groups	4074.47	60	67.91	

C. First Quarter Plane Geometry Course Grades

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	3132.25	64		
Between Means	74.44	4	18.61	.37
Within Groups	3057.81	60	50.96	

The time spent outside of class by the control group was only an estimate reported by the students and may have been exaggerated in some cases to impress the regular teacher or the experimenter. This fact should be considered when interpreting the results of the comparison between the control group and the experimental groups on the variables of time and efficiency.

The time out of class was added to the time in class to determine the total time spent on the unit by each student in the control group. The efficiency score for each member of this group was determined in the same way as for students in the experimental groups - the ratio of the post-test achievement score to the total time in minutes used to cover the material.

The means and standard deviations of the five treatment groups on each of the dependent variables are presented in Table 25.

An analysis of variance design was used to test the hypothesis of no differences between treatment groups on each of the dependent variables. The null hypothesis was tested with $\alpha = .05$. A summary of the results is contained in Table 26.

Hypothesis 7, viz., that no differences exist between the means of the experimental and control groups on achievement or retention, was accepted. However, the control group had a higher mean score on post-test achievement than any of the experimental groups. The differences between the control group and three of the experimental groups (1, 2 and 4) were sufficiently large that the probability of these differences occurring by chance under the null hypothesis was less than .10.

Table 25

Means of the Control Group and the Experimental Groups on the Dependent Variables

A. Post-test Achievement Scores

Statistic	Experimental Group				Control Group
	1	2	3	4	5
Mean	74.3	72.6	77.2	72.0	85.4
S.D.	17.8	17.5	19.5	18.5	16.5

B. Delayed Achievement Scores

Statistic	Experimental Group				Control Group
	1	2	3	4	5
Mean	72.1	78.3	80.3	69.8	79.8
S.D.	21.6	17.6	15.1	20.8	16.8

C. Retention Scores (in percent)

Statistic	Experimental Group				Control Group
	1	2	3	4	5
Mean	76.9	86.1	86.2	79.8	82.7
S.D.	15.2	11.0	8.9	13.9	9.6

D. Time Spent on Unit

Statistic	Experimental Group				Control Group
	1	2	3	4	5
Mean	442.3	439.6	479.4	511.5	941.0
S.D.	54.2	62.7	36.7	60.5	225.9

E. Efficiency Scores

Statistic	Experimental Group				Control Group
	1	2	3	4	5
Mean	16.9	17.3	16.0	14.1	9.6
S.D.	4.4	5.9	3.5	3.1	3.1

Table 26

Analysis of Variance Comparing the Control and Experimental Groups on the Dependent Variables

A. Post-test Achievement Scores

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	20991.94	64		
Between Means	1811.22	4	452.81	1.42
Within Groups	19180.72	60	319.68	

B. Delayed Achievement Scores

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	21518.86	64		
Between Means	1162.89	4	290.72	.86
Within Groups	20355.97	60	339.27	

C. Retention Scores

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	9113.14	64		
Between Means	774.44	4	193.61	1.39
Within Groups	8338.70	60	138.98	

Table 26 (con't)

Analysis of Variance Comparing the Control and Experimental Groups on the Dependent Variables

D. Time Spent on the Unit

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	3795262.46	64		
Between Means	2848349.38	4	712087.34	45.12**
Within Groups	946913.08	60	15781.89	

E. Efficiency Scores

Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F
Total	1579.39	64		
Between Means	589.42	4	147.35	8.93**
Within Groups	989.97	60	16.50	

**Significant. (probability less than .01)

Significant differences did exist between the control group and the experimental groups on the time spent on the programs and on mean efficiency scores. This was due to homework assignments given only to the control group.

CHAPTER V.

CONCLUSIONS AND SUMMARY

This study was designed to obtain empirical evidence on the relative effectiveness of two basic programing techniques in teaching a portion of high school geometry. Seven hypotheses were tested.

Conclusions Regarding Tests of the Hypotheses

A two way analysis of variance was used to test the null hypothesis of no differences between treatments or between ability levels. Hypotheses 1 and 3 were accepted; there were no significant differences between the experimental treatments in achievement or retention. However, both mixed-treatment groups had mean retention scores considerably higher than the two groups using one program exclusively. The use of a larger sample may have resulted in these differences being statistically significant.

Hypotheses 2 and 4 predicted no differences between ability levels in achievement and retention; both were rejected. The high ability students did significantly better than the low ability students in both achievement and retention. Two factors possibly contributed to these significant differences between the high and low ability students.

First, the nature of the material in geometry is such that verbal comprehension and verbal reasoning are very important. Both of these factors correlate highly with IQ scores. The programed unit required the students to work through 46 geometric proofs which demanded considerable concentration. Most of the proofs involved less than seven steps but still required the student to keep in mind a sizeable amount of information. This fact is not reflected in the error rate, especially on the linear program. The steps of the program were small enough that the student could usually answer the next question regarding a particular step in a proof without difficulty. Evidence supporting this point of view is the relatively low error rate of approximately 8% for the low ability students on the linear program. However, this does not mean that the student necessarily had a good grasp of the logical sequence or plan of the whole proof. This provides a real challenge to programers of material concerned with involved, logical arguments. Insuring that the student can take the next step successfully in a program, by sufficiently granulating the material and then arranging it systematically, is no guarantee that he will understand the logical development involved. Also, the student will not remember very well the facts he does learn if he fails to comprehend the logical structure and relationships of the concepts presented. It seems reasonable to assume that the logical relationships of the concepts developed in geometry are more difficult for the low ability student to comprehend by himself, even with the aid of programed materials, than for the high ability students.

The second factor which could possibly account for the differences in achievement and retention between the ability levels is the difficulty of the criterion test. A test which is so easy that practically all students answer over 80 or 90% of the questions correctly is a poor instrument for discrimination purposes. The post-test achievement scores in this study ranged from 35 to 103 out of a possible 108. The test was of sufficient difficulty and length that considerable room existed for individual differences to operate. Tests which are too short, too easy or too difficult are usually not desirable instruments for evaluating differences in treatments or levels in a research experiment.

The next two hypotheses involved the attitude questionnaire. Hypothesis 5 stated that the high ability students would show no preference for either the branch or the linear program. The prediction was made that the null hypothesis would be rejected, and the high ability students would indicate a strong preference for the more difficult branch program. The null hypothesis was rejected, but the preference was in the opposite direction of the one predicted. The high ability students in the mixed-treatment groups indicated a statistically significant preference (.05 level) for the linear program rather than the branch program.

One possible explanation for this surprising preference for the linear program is that the branch program may have been too difficult. The high ability students are accustomed to a high rate of success and this feeling was reinforced with the easier linear program. On the branch program, an error rate of 20 to 25% might not bother students who normally work at this level of success. However, an error rate of 15 or 20% on the branch program could be quite frustrating to the students who are accustomed to a much higher level of success.

Hypothesis 6 stated that the low ability students would indicate no preference for either the linear or the branch program. The prediction was made that the null hypothesis would be rejected, and the low ability students would show a strong preference for the easy, linear program. The low ability students in the mixed-treatment groups favored the linear program, as predicted, but the preference was not strong enough to be statistically significant at the .05 level. Therefore, Hypothesis 6 was accepted.

The combined mixed-treatment groups indicated a significant preference for the linear program. This finding that both mixed-treatment groups preferred the linear program supports Skinner's point of view that a high rate of positive reinforcement is self motivating and appeals to students of all levels of ability.

A reduction of the error rate of the branch program to 10 or 15%, by the inclusion of more steps, might result in greater achievement than with the present 20% error rate program. It would be interesting to see if a less difficult branch program also resulted in a more favorable attitude. Even if achievement remained the same, attitude might be more favorable, a prediction that would follow from the explanation given earlier for the finding that the higher ability students preferred the linear program.

Hypothesis 7 stated that no differences would exist between the control group and any of the experimental groups in achievement or retention. The null hypothesis was tested by analysis of variance at $\alpha = .05$. The hypothesis was accepted in regard to both achievement and retention. However, the mean post-test achievement score of the control group was considerably higher than any of the experimental groups. In fact, individual t-tests revealed that the differences between the control group and three of the experimental groups were sufficiently large that the differences could have occurred by chance under the null hypothesis with a probability less than .10.

It would be interesting to know how important the homework assignments were in accounting for the higher achievement scores of the control group. This could be experimentally determined in a future experiment by including a second control group which would cover the same unit of material but not have homework assignments.

The experimental groups did well on delayed achievement and retention when compared with the control group. Working on their own for two weeks without a teacher did not have adverse effects on remembering what they had learned over an intervening period of seven weeks. In fact, both mixed-treatment groups surpassed the control group on mean retention scores. One implication of this finding is that varying the type of program may aid retention.

Programs are being written in which constructed response and multiple-choice questions are freely intermixed. These may prove to be superior to either a pure linear program with constructed response questions or a pure branching program in which multiple-choice questions are used exclusively.

Comparison of These Findings with Related Studies

The analysis of variance revealed no significant differences between the experimental treatment groups in post-test achievement, delayed achievement or retention scores. These results are in agreement with the findings of Coulson and Silberman (4) who programed a portion of a college psychology course. They compared step size, mode of response and step sequencing separately and found no significant differences in achievement or retention. The present study combined the features of small step, constructed respnse and linear sequence into one program and still found no significant differences in achievement or retention when compared with a program combining the features of large step, multiple-choice response and branch sequencing.

These findings also agree with those of Silberman _et_ _al_ (17) who compared matched pairs of high school students on programed material in logic. The programs differed in having either a fixed or varied sequence of steps. The results indicated no significant differences in achievement.

We found that the high ability students did significantly better on achievement and retention than the low ability students. This agrees with the findings of Silberman _et_ _al_ (17) who reported that the poorer students did not do so well as the rest of the group following programed instruction in logic.

Shay (16) programed a unit on Roman numerals and divided fourth grade students into three ability levels on the basis of scores on the Henmon-Nelson Test of Mental Ability. He found significant differences between the ability levels on the total achievement test and on both the "rote" and "understanding" subtests. The high ability students did the best, followed by the average ability students, and finally the low ability students. This agrees with our findings that the high ability students are superior in achievement following programed instruction.

Shay was mainly interested in interaction between ability level and variations in step size. His covariance analysis revealed no significant interaction on achievement or time spent on the program materials. This agrees with our findings of no significant interaction between ability and treatment on the variables of achievement and time. We found no significant interaction between ability level and treatment on any of the variables tested. This result indicates that neither of the two programs compared in this study is more suitable for high or low ability students than the other program. Both programs resulted in a significant amount of learning for students of high and low ability with the high ability students doing considerably better on achievement and retention.

Our finding of little change in attitude toward programed instruction between questionnaires administered during the experiment and at the end agrees with those of Cassel and Ullom (2) (3) who found a high stability of attitudes toward programed instruction. They measured attitudes after one hour of programed instruction and at the end of the experiment using both average and high ability students.

Our finding that the favorable attitude toward programed instruction tended to disappear after the students were back in the regular classroom situation needs further exploration. More studies should be conducted in which attitudes are sampled after the subjects have been back in the normal routine for a period of time as well as while the experiment is being administered.

Our finding that the linear, small step, constructed response program took a considerably longer time to complete agrees with the findings of Coulson and Silberman (4) and Fry (11). The time factor in working through programed material may become a crucial variable in evaluating various programs. If studies continue to show that step size, mode of response and varied or fixed sequencing have little effect on achievement, then other variables must be considered. In this experiment, the average time spent on the branch program was one hour less than the average time spent on the linear program; yet the differences in achievement and retention were negligible. This difference in time required to cover the material in the two programs has practical implications. A savings of one hour in nine is a savings of approximately four weeks when projected over a full school year.

Another interesting finding was that the high ability students took as much time to cover the material as the low ability students. The largest difference was in Group 1 where only 12 minutes separated the mean times of the high and low ability subgroups.

One explanation for these small differences in time may have been that the students were told they would be graded on this unit of work the same way as other units in the course. So there was no reason for the better students to hurry through the material and get an average score on a test covering the material, knowing that it was only an experiment and would not influence their course grade. This factor should be considered when interpreting results of a learning experiment. Does the student consider himself merely a guinea pig working on experimental materials unrelated to the regular course work for the sole benefit of providing data for an experimenter? Or is the student ego involved, with the experiment covering an integral part of the regular course content and the student aware that he will be held responsible for the material on subsequent examinations? The latter case prevailed in this study and may account in part for the better students spending as much time as they did on the programs and then scoring significantly higher on the criterion test than the low ability students.

A promising avenue for research in programing has been introduced by Gagné and Paradise (12). They suggest that learning tasks be broken down into a hierarchy of subordinate learning sets. The degree to which these learning sets are easily recallable and the rate at which they are acquired seem to be more highly correlated with the final learning task than either general or specific mental abilities. Failure to achieve the desired learning can often be traced to subordinate learning sets left out of the program or not sufficiently integrated with other learning sets in the hierarchy. This approach seems especially appropriate in high school geometry where a knowledge of the axioms, postulates, defined terms and undefined terms in the axiomatic system are essential to the derivation of theorems later in the course.

Summary

This study was a comparison of a linear and a branch program covering a unit on parallel and perpendicular lines in high school plane geometry. A comparison was also made between programed instruction and regular classroom instruction with a teacher and traditional textbook.

The study involved 65 high school plane geometry students in three classes taught by the same teacher in Rantoul Township High School, Rantoul, Illinois. Two classes were divided into four treatment groups; two groups used the linear and branch programs exclusively and two groups switched from one program to the other halfway through the experiment. The Henmon-Nelson Test of Mental Ability (Grades 9-12, Form A) was used as a control variable to assign 48 students to four treatment groups and two ability levels by a stratified random sampling procedure. The third class with 17 students acted as a control group and covered the same unit of material with the regular teacher and textbook.

The relative effectiveness of the experimental treatment groups was judged on a) post-test achievement scores, b) delayed achievement scores, c) retention scores, d) attitude toward programed instruction, e) preference for the linear or the branch program, f) time spent on the material, and g) the efficiency in learning (a ratio of achievement to time spent on the materials).

A two way analysis of variance revealed no significant differences between the means of the four programed instruction groups on general mental ability, final freshman algebra grades or first quarter geometry grades. However, the students on the high ability level had mean scores superior to the low ability students on all three independent variables.

A two way analysis of variance revealed no significant differences between the means of the four groups on post-test achievement, delayed achievement or retention scores. However, on each of these variables, the students in the high ability subgroups did significantly better than the low ability students.

All four experimental groups had mean scores more favorable to programed instruction than regular classroom instruction midway through and at the completion of the experiment. Seven weeks later, the attitude questionnaire was administered a third time and the mean attitude scores of all four experimental groups had dropped to a neutral position, indicating no preference for either the programed instruction or the regular classroom instruction. No significant differences existed between the mean attitude scores of the four experimental groups or the ability subgroups at any of the three times the attitude questionnaire was administered.

The two mixed-treatment groups were asked to indicate their preference for the linear or the branching program. The total group and both high and low ability subgroups preferred the linear program. A t-test revealed that the mean preference scores of the total group and the high ability subgroup were significantly different from the neutral "no preference" score at the .05 level.

Group 4, which used the linear program exclusively, took significantly more time (.05 level) than Group 1 which used the branch program exclusively. Mean scores of efficiency were compared and Group 4, using the linear program, was the least efficient. The differences were not statistically significant.

The control group had a higher mean achievement score than any of the experimental groups, but the differences were not sufficiently large to reject the null hypothesis at the .05 level. Also, no significant differences were found between the control group and the experimental groups on delayed achievement or retention. The control group spent approximately the same amount of time in class as the experimental groups, but considerably more total time on the material due to homework assignments.

In conclusion, all five treatments resulted in a significant amount of learning during the two week experiment; in each treatment group the high ability students exceeded the low ability students in achievement and retention. The branch program was more efficient than the linear program timewise. Nevertheless, the students expressed attitudes more favorable to the linear program than the branch program.

BIBLIOGRAPHY

1. Beane,D. G. Parallel and Perpendicular Lines: by Programed Instruction. 2 Vols. Unpublished linear and branch programs. University of Illinois, 1961.

2. Cassel, R. N. and Ullom, W. L. "A Preliminary Evaluation of Programmed Instruction with Students of High Ability." Psychological Reports. 10, 1962, 223-228.

3. Cassel, R. N. and Ullom, W. L. "A Preliminary Evaluation of an Automatic Tutoring Machine." Journal of Secondary Education, 37, 1962, 117-124.

4. Coulson, J. E. and Silberman, H. F. "Effects of Three Variables in a Teaching Machine." Journal of Educational Psychology, 51, 1960, 135-143.

5. Crowder, N. A. "Intrinsically Programmed Teaching Devices." In Invitational Conference on Testing Problems. Proceedings. 1959, 40-52.

6. Crowder, N. A. The Arithmetic of Computers. Garden City, New York: Doubleday, 1960.

7. Crowder, N. A. and Martin, G. C. Adventures in Algebra. New York: Doubleday and Co., 1960.

8. Davis, F. B. Item-Analysis Data. Cambridge: Graduate School of Education, Harvard University, 1946.

9. Dixon, W. J. and Massey, F. J., Jr. Introduction to Statistical Analysis. New York: McGraw Hill Book Co., Inc., 1957.

10. Evans, J. L., Glaser, R., and Homme, L. E. "A Preliminary Investigation of Variation in the Properties of Verbal Learning Sequences of the 'Teaching Machine' Type." In Lumsdaine and Glaser (Eds.) Teaching Machines and Programmed Learning. Washington, D.C.: National Education Association, 1960, 446-451.

11. Fry, E. B. Teaching Machines: An Investigation of Constructed versus Multiple-Choice Response Modes. Unpublished Ph.D. Dissertation. University of Southern California, 1960.

12. Gagné, R. M. and Paradise, N. E. "Abilities and Learning Sets in Knowledge Acquisition." Psychological Monographs. Washington, D.C.: American Psychological Association, Vol. 75, No. 14, 1961.

13. Holland, J. G. and Skinner, B. F. The Analysis of Behavior. New York: McGraw Hill Book Co., 1961.

14. Lindquist, E. F. Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin Co., 1953.

15. Lumsdaine, A. A. and Glaser, R. (Eds.) Teaching Machines and Programmed Learning. Washington, D.C.: National Education Association, 1960.

16. Shay, C. B. "Relationship of Intelligence to Step Size on a Teaching Machine Program." Journal of Educational Psychology, 52, 1961, 98-103.

17. Silberman, H. F., Melaragno, R. J., Coulson, J. E., and Estavan, D. "Fixed Sequence Versus Branching Autoinstructional Methods." Journal of Educational Psychology, 52, 1961, 166-172.

18. Skinner, B. F. "Science of Learning and Art of Teaching." Harvard Educational Review, 24, 1954, 86-97.

19. Skinner, B. F. "Teaching Machines." Science, 128, 1958, 969-977.

20. Skinner, B. F. and Holland, J. G. "The Use of Teaching Machines in College Instruction." In Lumsdaine and Glaser (Eds.) Teaching Machines and Programmed Learning. Washington, D. C.: National Education Association, 1960, 159-172.

21. Stolurow, L. M. Teaching By Machine. Cooperative Research Monograph No. 6. U. S. Department of Health, Education and Welfare. Office of Education. Washington, D.C.: Government Printing Office, 1961.

22. Walker, C. C. A Comparison of Overt and Covert Response in Programed Learning. Unpublished MA thesis. University of Illinois, 1961.

23. Walker, Helen M. and Lev, J. Statistical Inference. New York: Holt and Co., 1953.

24. Welchons, A. M. and Krickenberger, W. R. New Plane Geometry. Boston: Ginn and Co., 1956.

APPENDIX A

CRITERION ACHIEVEMENT TEST

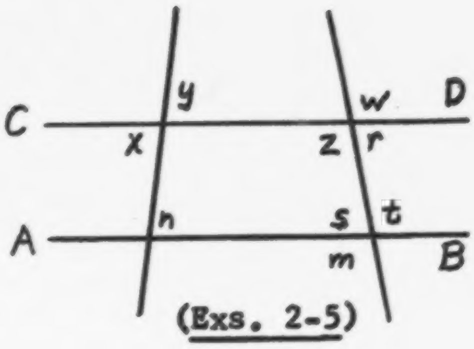
Test on Parallel and Perpendicular Lines

Name ______________________

____

Part I. True-False Statements
(Write True or False in blank)

1. Through a point outside a line there can be two lines parallel to the given line. ______

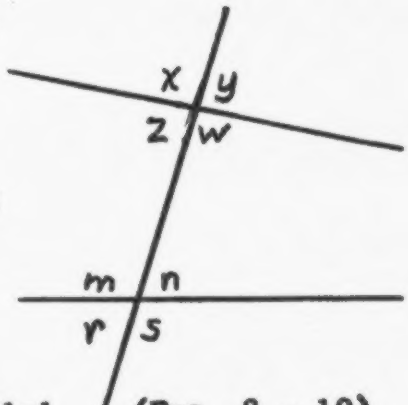


(Exs. 2-5)

2. If $\angle z + \angle s = 180^\circ$, $AB \parallel CD$. ______
3. If $AB \parallel CD$, $\angle x = \angle t$. ______
4. If $\angle w = \angle t$, $AB \parallel CD$. ______
5. If $AB \parallel CD$, $\angle n + \angle s = 180^\circ$. ______

Part II. Completing Statements.

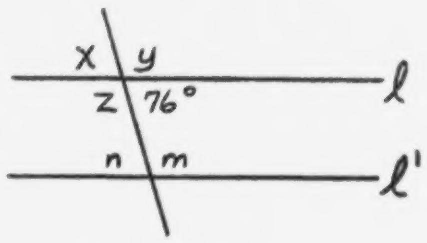
6. Two lines parallel to a third line are ______.
7. Two lines perpendicular to a third line are ______.



(Exs. 8 - 10)

8. $\angle x$ and $\angle$ ______ form a pair of vertical angles.
9. $\angle w$ and $\angle$ ______ form a pair of corresponding angles.
10. $\angle z$ and $\angle$ ______ form a pair of alternate interior angles.
11. If a line is perpendicular to one of two parallel lines, it is ______ to the other.
12. The two acute angles of a right triangle are ______.

In the figure, ℓ and ℓ' are parallel.



(Exs. 13-15)

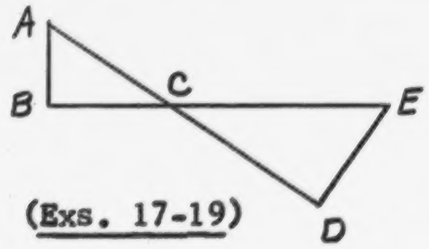
13. ∠ z = ________ degrees.

14. ∠ m = ________ degrees.

15. ∠ n = ________ degrees.

16. If one acute angle of a right triangle is 30°, the side opposite this angle is ________ the hypotenuse.

Given AD ⊥ DE, AB ⊥ BE, and ACD and BCE stright lines.



(Exs. 17-19)

17. ∠ B = ∠ ________

18. ∠ ACB = ∠ ________

19. ∠ A = ∠ ________

20. One angle of an equilateral triangle contains ________ degrees.

Part III. Applications

21. If the acute angles of a right triangle are 40° and 50° respectively, how many degrees are there in the angle formed by their bisectors? ________.

22. How long is the hypotenuse of a 30°-60° right triangle if the side opposite the 30° angle is 4 inches? ________.

23. In △ ABC, ∠ A = 42° and ∠ B = 70°. How many degrees are there in ∠ C? ________.

24. In △ DEF, ∠ D = 42° and ∠ E = 23°. How many degrees are there in either exterior angle at F? ________.

25. The vertex angle of an isosceles triangle is 66°. How many degrees are there in each base angle? ________.

Part IV. Supplying Reasons

26. In △s ABC and DEF, ∠ A = ∠ D, and ∠ B = ∠ E. Why does ∠ C = ∠ F?

____

____

27. If $a \parallel b$, $\angle o = \angle s$. Why? ________

____

____

28. If $c \parallel d$, $\angle w + \angle o = 180^\circ$. Why? ___

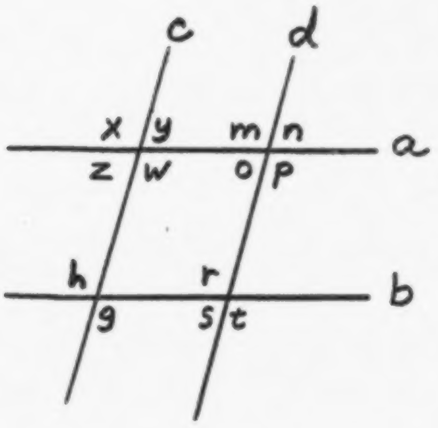
____

____

29. If $\angle g = \angle r$, $c \parallel d$. Why? ________

____

____



(Exs. 27-30)

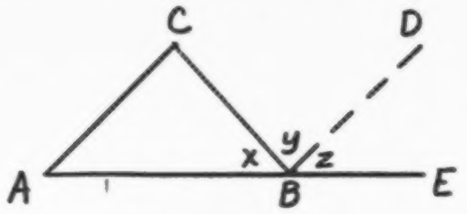
30. If $\angle z$ and $\angle h$ are supplementary, $a \parallel b$.

Why? ____________________________

____

Given $\triangle ABC$ with $BD \parallel AC$.

To prove that $\angle ABC + \angle C + \angle A = 1$ st. $\angle$.



(Exs. 31-35)

Proof: Statements	Reasons
31. $AC \parallel DB$	31.
32. $\angle x + \angle y + \angle z = 1$ st. $\angle$.	32.
33. $\angle A = \angle z$.	33.
34. $\angle C = \angle y$.	34.
35. $\angle ABC + \angle C + \angle A = 1$ st. $\angle$.	35.

Part V. Drawing a Figure for a Theorem and Stating the Hypothesis and Conclusion

Draw the figure, label it, and state what is given and what is to be proved in terms of the figure.

36. If a perpendicular drawn from one vertex of a triangle bisects the opposite side, it bisects the angle at the vertex.

Given: 1. ______________________ **Figure**

2. ______________________

3. ______________________

Prove: ______________________

37. If the median of a triangle is equal to one half the side to which it is drawn, the triangle is a right triangle.

Given: 1. ______________________ **Figure**

2. ______________________

3. ______________________

Prove: ______________________

38. The bisector of the vertex angle of an isosceles triangle is perpendicular to the base.

Given: 1. ______________________ **Figure**

2. ______________________

Prove: ______________________

39. The exterior angle at the base of an isosceles triangle is equal to the angle formed by the bisectors of the base angles.

Given: 1.________________________ **Figure**

2. ________________________

3. ________________________

Prove: ________________________

40. If in $\triangle$ ABC perpendiculars from A and B to the opposite sides are equal, the triangle is isosceles.

Given: 1. ________________________ **Figure**

2. ________________________

3. ________________________

Prove: ________________________

Part VI. Converses and Inverses

41. What parts of a theorem are interchanged when its complete converse is written? ____________________ and ____________________.

42. Write the converse of: If an altitude of a triangle bisects one side, the triangle is isosceles.________________________________

____

____

43. Write the converse of: If two angles of a triangle are equal, the sides opposite these angles are equal. _________________________

____

____

44. Write the inverse of: If a line through the vertex of an isosceles triangle bisects the base, it is perpendicular to the base.

____

____

45. Write the inverse of: If two lines form equal corresponding angles with a transversal, the lines are parallel. ______________

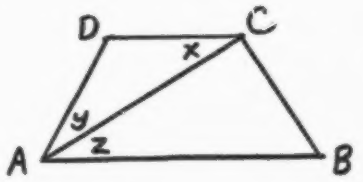
____

____

Part VII. Proofs

Given figure ABCD with $AB \parallel DC$ and $AD = DC = CB$.

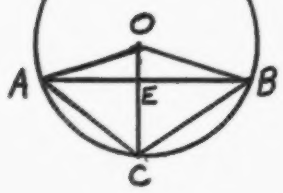
To prove that AC bisects ∠BAD.



Proof:

Statements	Reasons
46.	46.
47.	47.
48.	48.
49.	49.
50.	50.

Given circle 0 with radii OA and OB, AC = CB, and OC and AB intersecting at E.



Prove that AE = EB.

Proof: Statements	Reasons
51.	51.
52.	52.
53.	53.
54.	54.

APPENDIX B

ATTITUDE QUESTIONNAIRE

Name ______________________

Questionnaire

1. I want to continue one more week using these materials. (Circle one)

No definitely		It doesn't matter		Yes, very much
1	2	3	4	5

2. I want to continue the rest of this semester in geometry using programed materials. (Circle one).

No definitely		It doesn't matter		Yes, very much
1	2	3	4	5

3. (Check one) The use of programed materials:

a. _____ is an excellent method of teaching.
b. _____ is a good method of teaching.
c. _____ is an acceptable method of teaching.
d. _____ is an undesirable method of teaching.
e. _____ is a very poor method of teaching.

4. In comparing programed instruction with a teacher-and-textbook, I believe in plane geometry I learned: (Check one)

a. _____ much more using programed materials on my own.
b. _____ a little more using programed materials on my own.
c. _____ about the same either way
d. _____ a little more with teacher-and-textbook in a regular class.
e. _____ much more with teacher-and-textbook in a regular class.

5. What was your reaction to the procedure in programed instruction of telling you immediately whether your answers were right or wrong? (Check one)

a. _____ I thought it was one of the best features of programed instruction.
b. _____ I liked it better than the usual waiting until the next day to check homework problems.
c. _____ It doesn't matter to me one way or the other.
d. _____ I would rather work a whole set of problems and then be told the right answers.
e. _____ I would much rather have the opportunity to work through a whole group of problems on my own without being told every step of the way if I am right or not.

6. If I had my choice in a math course, I would prefer to be in a class: (Check one)

a. _____ Using programed materials all the time.
b. _____ Using programed materials most of the time.
c. _____ Use programed materials half the time and have a teacher-and-textbook half the time.
d. _____ Have a teacher-and-textbook most of the time.
e. _____ Have a teacher-and-textbook all the time.

7. My general impression of programed instruction is: (Check one)

a. _____ I like it very much.
b. _____ I like it fairly well.
c. _____ My impression is neither favorable nor unfavorable.
d. _____ I do not like it very well.
e. _____ I dislike it very much.

Comparison of the Multiple-Choice vs. the Completion Question Programs

8. Next week in studying plane geometry, I want to: (Check one)

a. _____ use the multiple-choice program.
b. _____ use the completion question program
c. _____ go back to the usual procedure of having a teacher and a regular textbook.

9. For the rest of this semester in plane geometry I want to: (Indicate 3 choices: 1st choice: 1; 2nd choice: 2; 3rd choice: 3.)

a. _____ use the multiple-choice program all the time.
b. _____ use the completion type program all the time.
c. _____ be in a class with a teacher and use a regular textbook.
d. _____ use programed materials but alternate between the multiple-choice and completion programs.
e. _____ have a teacher and regular textbook part of the time and use a multiple-choice program on my own part of the time.
f. _____ have a teacher and regular textbook part of the time and use a completion question program part of the time.
g. _____ alternate all three types of instruction: 1) teacher-and-textbook; 2) multiple-choice program; and 3) completion program.
h. _____ other choice. Describe: ______________________________
____

10. In question 9 above, indicate with a letter "L" the choice you would least prefer.

11. What I like most about programed instruction is: ____________________

____

12. What I like least about programed instruction is: ____________________

____

13. The percentage of material I would estimate that I have learned from this program is (Circle number which is closest to the value of your estimate.)

0 10 20 30 40 50 60 70 80 90 100

APPENDIX C

TIME SHEET

Name: ____________________

Time spent on plane geometry outside of class. Indicate time in minutes (such as 5, 10, 15, or 30 minutes).

Monday ____________________

Tuesday ____________________

Wednesday ____________________

Thursday ____________________

Friday ____________________

Saturday ____________________

Sunday ____________________

Turn in this sheet each Monday before class.

APPENDIX D

DATA FOR INDIVIDUAL SUBJECTS

Group 1

High Ability Subgroup	Deviation IQ Score	Final Algebra Grade	1st Quarter Geometry Grade	Achievement Post-Test Score	Delayed Achievement Score	Retention Score %
S 1	127	86	95	76	90	88
S 2	120	94	94	89	95	96
S 3	120	97	99	99	104	98
S 4	119	86	96	98	88	85
S 5	116	88	87	61	55	64
S 6	115	84	87	79	67	65
Low Ability Subgroup						
S 7	114	82	87	83	85	92
S 8	109	74	87	59	70	80
S 9	105	71	75	74	72	76
S 10	98	74	83	48	44	58
S 11	95	95	85	80	64	68
S 12	92	72	72	45	31	53

High Ability Subgroup	Attitude Score Time-1	Attitude Score Time-2	Attitude Score Time-3	Preference Score (Lin.vs Br.)	Total Time (Min.)	Efficiency score	Error Rate (Branch)
S 1	31	33	18	22	326	.23	20
S 2	31	30	24	17	540	.16	10
S 3	26	24	24	21	432	.23	03
S 4	35	33	32	20	476	.21	16
S 5	32	21	16	21	438	.14	30
S 6	27	26	20	21	406	.19	10
Low Ability Subgroup							
S 7	29	26	21	21	471	.18	37
S 8	26	20	19	20	383	.15	10
S 9	26	24	23	21	470	.16	33
S 10	17	14	11	22	433	.11	33
S 11	30	29	21	24	452	.18	26
S 12	29	32	24	24	480	.09	30

Group 2

High Ability Subgroup	Deviation IQ Score	Final Algebra Grade	1st Quarter Geometry Grade	Achievement Post-Test Score	Delayed Achievement Score	Retention Score %
S 13	138	90	97	90	97	96
S 14	136	96	97	100	105	98
S 15	124	89	87	68	78	85
S 16	120	84	84	80	78	86
S 17	117	90	93	81	80	88
S 18	115	88	90	65	77	83
Low Ability Subgroup						
S 19	114	83	80	51	67	92
S 20	114	81	87	69	73	84
S 21	112	96	94	99	100	93
S 22	107	93	87	57	78	95
S 23	105	90	92	63	69	75
S 24	101	70	87	48	38	58

High Ability Subgroup	Attit. Score Time-1	Attit. Score Time-2	Attit. Score Time-3	Preference Score (Lin.vs Br.)	Total Time (Min.)	Efficiency Score	Error Rate (Lin.)	Error Rate (Br.)
S 13	34	33	24	24	366	.25	07	17
S 14	26	28	26	17	357	.28	05	07
S 15	19	25	15	11	377	.18	05	06
S 16	21	24	15	15	516	.16	08	33
S 17	17	20	16	13	552	.15	05	22
S 18	32	31	26	12	466	.14	02	18
Low Ability Subgroup								
S 19	32	27	17	16	434	.12	02	06
S 20	25	22	16	21	438	.16	02	17
S 21	30	32	26	24	374	.26	06	13
S 22	21	20	26	23	491	.12	04	17
S 23	9	9	14	22	434	.15	14	36
S 24	33	30	30	16	470	.10	09	19

Group 3

High Ability Subgroup	Deviation IQ Score	Final Algebra Grade	1st Quarter Geometry Grade	Achievement Post-Test Score	Delayed Achievement Score	Retention Score %
S 25	147	80	96	103	97	92
S 26	126	90	94	95	96	96
S 27	120	92	89	90	84	84
S 28	117	97	93	93	91	86
S 29	116	90	83	66	72	84
S 30	115	79	94	68	82	93
Low Ability Subgroup						
S 31	113	95	97	63	78	89
S 32	113	90	89	100	101	96
S 33	111	90	84	73	69	70
S 34	108	91	92	54	78	93
S 35	105	83	78	41	48	80
S 36	101	91	87	80	67	71

High Ability Subgroup	Attit. Score Time-1	Attit. Score Time-2	Attit. Score Time-3	Preference Score (Lin. vs Br.)	Total Time (Min.)	Effi-ciency Score	Error Rate (Lin.)	Error Rate (Br.)
S 25	24	17	15	15	507	.20	05	17
S 26	28	26	21	24	466	.20	01	15
S 27	15	14	16	21	523	.17	09	22
S 28	25	22	19	11	481	.19	03	06
S 29	33	32	17	10	480	.14	07	20
S 30	33	34	18	11	398	.17	07	29
Low Ability Subgroup								
S 31	16	15	13	16	477	.13	02	07
S 32	24	14	28	17	509	.20	09	25
S 33	25	22	18	24	503	.15	04	08
S 34	21	25	18	16	461	.12	08	28
S 35	30	31	32	10	432	.09	25	34
S 36	18	21	18	21	516	.16	15	33

Group 4

High Ability Subgroup	Deviation IQ Score	Final Algebra Grade	1st Quarter Geometry Grade	Achievement Post-Test Score	Delayed Achievement Score	Retention Score %
S 37	137	89	95	97	95	92
S 38	126	86	93	87	81	76
S 39	119	92	87	68	61	79
S 40	118	98	98	98	108	100
S 41	115	80	75	61	55	74
S 42	114	91	92	80	81	91
Low Ability Subgroup						
S 43	113	93	96	89	89	91
S 44	109	83	88	62	65	84
S 45	108	75	82	65	52	66
S 46	104	88	89	59	62	83
S 47	103	65	70	63	39	48
S 48	90	72	80	35	50	74

High Ability Subgroup	Attit. Score Time-1	Attit. Score Time-2	Attit. Score Time-3	Preference Score (Lin. vs Br.)	Total Time (Min.)	Effici-ency Score	Error Rate (Linear)
S 37	20	16	19	22	543	.18	07
S 38	28	26	25	11	583	.15	05
S 39	20	25	23	20	436	.16	17
S 40	33	33	33	16	511	.19	05
S 41	30	30	19	11	490	.12	05
S 42	17	21	13	19	498	.16	11
Low Ability Subgroup							
S 43	32	31	32	11	614	.14	04
S 44	21	17	17	23	588	.11	11
S 45	33	31	25	20	455	.15	06
S 46	23	20	16	19	501	.12	09
S 47	35	35	33	16	500	.13	11
S 48	17	16	11	22	429	.08	07

Control Group 5

High Ability Subgroup	Deviation IQ Score	Final Algebra Grade	1st Quarter Geometry Grade	Achievement Post-Test Score	Delayed Achievement Score	Retention Score %
S 49	126	79	88	92	72	77
S 50	125	99	97	102	105	97
S 51	122	93	97	89	70	72
S 52	116	77	80	85	84	89
S 53	114	96	96	102	99	91
S 54	113	88	98	103	98	90
S 55	113	85	88	96	86	85
S 56	111	97	96	104	104	96
S 57	111	85	82	69	57	62
Low Ability Subgroup						
S 58	109	80	90	86	92	93
S 59	108	98	92	95	88	89
S 60	108	93	87	86	80	80
S 61	105	91	89	83	73	78
S 62	103	77	90	88	76	81
S 63	100	75	87	49	51	76
S 64	95	87	80	67	67	75
S 65	88	73	72	55	55	75

High Ability Subgroup	Total Time (Min.)	Efficiency Score
S 49	760	.12
S 50	680	.15
S 51	910	.10
S 52	925	.09
S 53	931	.11
S 54	840	.12
S 55	860	.11
S 56	905	.11
S 57	965	.07
Low Ability Subgroup		
S 58	732	.12
S 59	1275	.07
S 60	668	.13
S 61	930	.09
S 62	768	.11
S 63	1068	.05
S 64	1480	.05
S 65	1300	.04

APPENDIX E

DISTRIBUTION OF RESPONSES TO INDIVIDUAL ITEMS ON THE ATTITUDE QUESTIONNAIRE (SECTION 1)*

		Time 1		Time 2		Time 3	
		No.	%	No.	%	No.	%
Item 1							
Response:	5	25	52	18	38	5	10
	4	5	10	6	12	9	19
	3	16	34	17	35	15	32
	2	1	2	2	4	9	19
	1	1	2	5	10	10	20
Item 2							
Response:	5	18	38	11	23	4	8
	4	3	6	7	15	7	15
	3	14	29	14	29	7	15
	2	3	6	3	6	10	20
	1	10	20	13	27	20	42
Item 3							
Response:	a	10	20	7	15	4	8
	b	22	46	26	54	15	32
	c	14	29	11	23	24	50
	d	1	2	3	6	5	10
	e	1	2	1	2	0	0
Item 4							
Response:	a	8	16	7	15	2	4
	b	12	25	10	20	5	10
	c	8	16	12	25	9	19
	d	14	29	14	29	20	42
	e	6	12	5	10	12	25
Item 5							
Response:	a	15	32	24	50	25	52
	b	28	58	18	38	16	34
	c	3	6	2	4	3	6
	d	1	2	3	6	3	6
	e	1	2	1	2	1	2

*The wording of the items and the alternatives is given in APPENDIX B.

		Time 1		Time 2		Time 3	
		No.	%	No.	%	No.	%
Item 6							
Response:	a	9	19	6	12	1	2
	b	12	25	14	29	6	12
	c	15	32	12	25	15	32
	d	10	20	15	32	18	38
	e	2	4	1	2	8	16
Item 7							
Response:	a	21	44	16	34	6	12
	b	15	32	18	38	19	40
	c	6	12	8	16	14	29
	d	6	12	6	12	9	19
	e	0	0	0	0	0	0